Análisis clásico elemental

Segunda edición

Análisis clásico elemental

Segunda edición

Jerrold E. Marsden
University of California, Berkeley

Michael J. Hoffman
California State University, Los Ángeles

Versión en español

Óscar Alfredo Palmas Velasco
Universidad Nacional Autónoma de México

Colaboración técnica

José Antonio Cuesta Ruiz
Universidad Carlos III
Madrid, España

Addison-Wesley Iberoamericana

Argentina • Chile • Colombia • España
Estados Unidos • México
Perú • Puerto Rico • Venezuela



Versión en español de la obra *Elementary Classical Analysis, 2nd. ed.* de Jerrold E. Marsden y Michael J. Hoffman, publicada originalmente en Estados Unidos por W.H. Freeman and Company, Nueva York, Nueva York y Oxford © 1993. Reservados todos los derechos.

First published in the United States by W.H. Freeman and Company, New York, New York and Oxford. Copyright 1993. All rights reserved.

Esta edición en español es la única autorizada.

ADDISON-WESLEY IBEROAMERICANA
Malabia 2363-2°G, Buenos Aires 1425, Argentina
Cruz 1469 Depto 21 Independencia, Santiago, Chile
Apartado Aéreo 241-943 Santafé de Bogotá, Colombia
Espalter 3 bajo, Madrid 28014, España
1 Jacob Way, Reading, Mass. 01867, E.U.A.
Apartado Postal 22-012, México D.F. 14000, México
Jr. San Antonio Este, núm. 658, depto. D, Urb. Ventura Rossi, Lima 25, Perú
El Monte Mall 2° piso, oficina 19-B, Ave. Muñoz Rivera Hato Rey, 00918, Puerto Rico
Apartado Postal 51454, Caracas 1050-A, Venezuela

Impreso en Estados Unidos. *Printed in U.S.A.*

ISBN 0-201-65369-9

1 2 3 4 5 6 7 8 9 10-MA-01 00 99 98 97

Contenido

Prefacio

Esta segunda edición tiene el mismo objetivo básico que la primera: presentar el análisis clásico elemental en un marco concreto, enfatizando las técnicas específicas importantes para el análisis clásico y sus aplicaciones. Como la primera edición, este libro se limita al análisis que usa el sistema de los números reales sin el cálculo vectorial o el análisis complejo. El lector puede revisar estos temas en *Vector Calculus*, 3a. edición, de Marsden y Tromba (W.H. Freeman and Company, 1987)[1] y *Basic Complex Analysis*, 2a. edición, de Marsden y Hoffman (W.H. Freeman and Company, 1987). El material de este libro está normalmente orientado a estudiantes de tercer o cuarto año de licenciatura, que han recibido ya un curso de cálculo vectorial o cálculo de varias variables y un curso de álgebra lineal.

Mantenemos el marco de referencia lo más concreto posible, presentando el análisis en el espacio euclídeo la mayor parte del tiempo, sin descuidar las aplicaciones. De hecho, cada vez más estudiantes de ingeniería y ciencias físicas siguen este tipo de curso y este libro pretende cubrir sus necesidades, al igual que las del estudiante de matemáticas puras. Por ejemplo, no dudamos en presentar un modelo extraído de la teoría de control o de la mecánica cuántica cuando ello es oportuno. Pensamos que este equilibrio entre los temas puros y los aplicados es importante para los estudiantes, tanto de matemáticas puras como aplicadas, en particular en esta época en que la frontera entre las dos áreas es cada vez más difusa.

Esta edición contiene más material de espacios métricos que la primera, sobre todo en aras de una exposición eficiente. Los nuevos temas de esta área están esparcidos a lo largo de los primeros cinco capítulos. Un segundo cambio fundamental respecto a la primera edición es la inclusión de nuevo material acerca del cálculo de una variable, sobre todo al final del capítulo 4. Varios lectores pensaron que sería útil un mayor repaso de este material, pues muchos estudiantes no han tenido un curso básico de cálculo con un fundamento teórico sólido. Estamos de acuerdo con ellos; sin embargo, damos por hecho que el estudiante ha recibido algún curso de cálculo de varias variables (o, al menos, tiene nociones sobre las derivadas parciales) y algún curso de álgebra

[1] Versión en castellano: J. Marsden y A. Tromba, *Cálculo vectorial*. 3ª ed., Addison-Wesley Iberoamericana, 1991. (*N. del R.T.*)

lineal (por lo menos conoce las matrices y las transformaciones lineales). Un tercer cambio fundamental es que hemos escrito de nuevo gran parte del primer capítulo acerca de los fundamentos del sistema de los números reales. Este tema está ahora más detallado que en la primera edición. Hemos decidido conservar como el axioma básico de la completitud la convergencia de sucesiones crecientes monótonas, ya que éste es el concepto más intuitivo y transparente. Por supuesto, pasamos rápidamente a la propiedad del supremo y la convergencia de sucesiones de Cauchy.

También hemos querido conservar el formato de la primera edición, que proporciona las demostraciones técnicas al final de cada capítulo pero presenta las ideas clave en el texto. Parece que esto ha sido bien recibido por la mayoría de los lectores de la primera edición y seguimos pensando que constituye un recurso pedagógico adecuado para un curso como éste. No es que con ello se pretenda rehuir de alguna manera las demostraciones; por el contrario, el propósito es dar dos enfoques de la demostración: uno de la forma en que los matemáticos conciben las demostraciones (los secretos profesionales, por así decirlo) y el otro de la forma en que los matemáticos escriben las demostraciones formales.

Durante la preparación de esta nueva edición, nuestros estudiantes nos han proporcionado muchos comentarios valiosos, por los cuales les estamos muy agradecidos. También queremos agradecer a Susan Knapp, Diana Siemens y Archetype Publishing Inc. su excelente trabajo en la producción de este libro.

Jerrold E. Marsden

Michael J. Hoffman

Prefacio a la primera edición

Este libro está diseñado para un curso de dos trimestres o para un curso de uno o dos semestres de cálculo avanzado e introducción al análisis real. El libro es clásico en el sentido de que trata el cálculo y las series de Fourier en el espacio euclídeo. Sólo hacemos unas cuantas referencias a temas "modernos", como la integración de Lebesgue, las distribuciones y la mecánica cuántica. Hemos resistido a la tentación de incluir el análisis vectorial (el teorema de Stokes y temas relacionados). En la mayoría de las carreras este tema aparece antes en el segundo año a un nivel más informal (véase, por ejemplo, J. Marsden y A. Tromba, *Cálculo vectorial, op. cit.*) y tal vez después en el contexto de la teoría de variedades para los estudiantes que tengan tales inclinaciones.

Al presentar el material, hemos sido deliberadamente concretos, con la mira hacia una comprensión sólida del caso euclídeo y presentando la abstracción sólo mediante ejemplos. Es decir, si se comprenden adecuadamente los espacios euclídeos, basta un pequeño salto para entender otros espacios, como el espacio de las funciones continuas y los espacios métricos abstractos. En el contexto del espacio de las funciones continuas podemos ver el poder de los métodos de los espacios métricos abstractos. Cuando la teoría general se presenta demasiado pronto, el estudiante se confunde acerca de su importancia; en consecuencia, puede llegar a desperdiciarse mucho tiempo de enseñanza.

El libro presupone que el lector conoce algo de cálculo; es decir, que sabe derivar e integrar las funciones usuales. En un sentido estricto, la teoría se desarrolla lógicamente y requiere pocos requisitos previos, aunque se necesita conocer el cálculo para comprender los ejemplos y ejercicios. Además, es recomendable pero no esencial cierto contacto con las derivadas parciales y las integrales múltiples. El capítulo 6, de diferenciación, necesita los fundamentos del álgebra lineal; específicamente, el estudiante debe saber lo que es una transformación lineal y la matriz que la representa.

Cada capítulo se organiza como sigue. Hay varias secciones que contienen las definiciones, enunciados de los teoremas, ejemplos y problemas sencillos. Una vez que el estudiante domine los teoremas y pueda resolver los problemas sencillos, puede pasar al final de cada capítulo para dominar las demostraciones técnicas. En esta parte se dan más ejemplos y ejercicios. Los ejercicios más sencillos después de cada sección permiten al estudiante dominar el material conforme éste avanza. Los ejercicios al final del

capítulo necesitan entonces, con frecuencia, un conocimiento integrado de todo el capítulo o de los capítulos anteriores, incluyendo las demostraciones de los teoremas. Este plan ha funcionado en las clases. Cuando las clases se dedican a la explicación de los teoremas y se dan solamente algunas demostraciones, es mucho más fácil para el estudiante ver lo que ocurre. Hemos visto que con el uso de este enfoque, en cada clase se cubren una o dos secciones.

El capítulo introductorio contiene la terminología esencial. El estudiante interesado en las complejidades de la teoría de conjuntos puede consultar el apéndice, amablemente proporcionado por el profesor I. Fáry.

El capítulo 1 contiene material acerca de la estructura básica de la recta real, necesario para los desarrollos posteriores. Utilizamos un poco de tiempo para los axiomas algebraicos y nos concentramos en la propiedad de completitud. Los axiomas algebraicos se ven por lo general en los cursos básicos de álgebra; como el estudiante está acostumbrado a trabajar con los números reales, parece lógico aceptar como válidas las habilidades algebraicas básicas.

Los capítulos 2 y 3 tratan la topología de $\mathbb{R}^n$ utilizando sólo la estructura métrica básica de $\mathbb{R}^n$. Esto se hace para que la transición a otros espacios métricos, como el espacio de las funciones continuas, que se trata posteriormente, sea casi automática.

Se evita aquí una introducción completa y prematura a los espacios métricos abstractos. La experiencia ha demostrado que se necesitan casi dos semanas más en este nivel para lograr esta abstracción, ya que uno tiene que pasar por los "extraños" espacios métricos usuales que confunden a los estudiantes. El tiempo ahorrado se puede utilizar posteriormente para temas más útiles como el teorema de Ascoli, el teorema de Stone-Weierstrass, los teoremas de punto fijo y las ecuaciones diferenciales o integrales.

El capítulo 4 continúa el desarrollo, tratando los resultados básicos de continuidad. El capítulo 5 da las propiedades más finas de las funciones continuas relacionadas con la convergencia uniforme. En las secciones 5.5–5.9 se presentan varios temas especializados, de los que puede hacerse una selección.

El capítulo 6 trata la diferenciación, haciendo un cierto uso del álgebra lineal, y abarca todos los temas habituales del cálculo diferencial para funciones de varias variables. Se da un tratamiento bastante completo de los máximos y los mínimos, con un análisis opcional del lema de Morse en el capítulo 7. Ese capítulo 7 tiene como tema principal un análisis completo de los teoremas de la función inversa y de la función implícita. También se dan los teoremas de existencia para ecuaciones diferenciales ordinarias y los multiplicadores de Lagrange (extremos condicionados).

El capítulo 8 trata los fundamentos de la integración. Tal vez algunos profesores quieran enseñar este material antes del capítulo 7. En este capítulo tratamos la integral de Riemann, pero incluimos el teorema de Lebesgue y los conjuntos de medida nula. Una sección opcional revisa rápidamente las distribuciones, ilustradas por la función δ.

El siguiente capítulo demuestra los dos teoremas fundamentales relativos a las integrales múltiples: la reducción a integrales iteradas y la fórmula para el cambio de variables, con numerosas aplicaciones.

El último capítulo, el 10, es un tratamiento bastante completo de las series de Fourier, desde el punto de vista de los espacios con producto interno. Algunos temas como éste son útiles para los estudiantes de los cursos de introducción al análisis, ya que van más allá de la simple "rigorización" de los temas que ya conocen. Una característica poco común de nuestra presentación es la inclusión de algunas aplicaciones a las ecuaciones diferenciales y la mecánica cuántica.

Por supuesto, los profesores tienen diferentes opiniones en cuanto al rigor, el papel de la intuición, la elección de temas, etcétera. Tal vez sean útiles unas cuantas observaciones acerca de las variantes en la forma de presentación del material de este libro para aquellos que deseen adaptarlo a su estilo personal.

En primer lugar, en los capítulos del 2 al 4, es posible enfatizar los espacios métricos abstractos sin cambiar materialmente el texto. De hecho, es un buen ejercicio para los estudiantes que hagan esta adaptación, ya que una vez que conocen la demostración "correcta" en $\mathbb{R}^n$, es bastante divertido y provechoso generalizarla. A este respecto, al final del capítulo 5 aparece una tabla, proporcionada por R. Gulliver, que indica los teoremas válidos para espacios métricos generales.

Parte del material del capítulo 5 es un poco más avanzado y puede posponerse. Además, si se desea un desarrollo lógico completo, la derivación y la integración de funciones de una variable deben preceder al capítulo 5; esto depende de la formación de los estudiantes. Este material, en sus aspectos más básicos, se usa en las secciones 5.3, 5.6 y 7.5. En la práctica, hemos visto que el uso de algunas técnicas de cálculo antes de su presentación "correcta" en el curso sólo produce rechazo en los mejores estudiantes. Esto nos parece saludable, pero algunos podrían preferir cambiar el orden de la presentación.

Al principio del capítulo 6, conviene repasar un poco de álgebra lineal; en particular, la definición de la matriz de una transformación lineal. También es una buena ocasión para revisar el ejemplo 4 al final del capítulo 4.

Para un curso de un semestre, hay que recortar algunos temas para llegar al capítulo 10 (tales como las secciones 5.5–5.9 y 7.3–7.7). En un curso de dos trimestres, se puede cubrir todo el texto (tal vez omitiendo las secciones 5.8, 5.9, 7.3, 7.4, 7.6, 7.7, 10.7 y 10.8).

Los símbolos utilizados en este texto son los comunes, excepto tal vez los siguientes: $\mathbb{R}$ denota la recta numérica real, $\mathbb{C}$ denota los números complejos, $\mathbb{R}^n$ denota el espacio euclídeo n-dimensional, "sii" quiere decir "si y sólo si" y ▮ denota el final de una demostración. La notación $]a, b[$ se usa para indicar el intervalo abierto que consta de todos los números reales x que satisfacen $a < x < b$. Esta notación europea evita la confusión con el par ordenado (a, b). La notación $x \mapsto f(x)$ indica que x se asocia con $f(x)$ mediante f. La notación $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ indica que f aplica el dominio A en $\mathbb{R}^m$. En ciertas ocasiones, $\Rightarrow$ se usa para denotar "implica". El símbolo $A \setminus B$ denota los elementos del conjunto A que no son de B y $x \in A$ indica que x es un elemento de A.

Las secciones, los teoremas y las definiciones se numeran de manera consecutiva dentro de cada capítulo. Una referencia como "teorema 24" o "ejercicio 3" se aplica al material dentro del propio capítulo o sección; en caso contrario, se cita el número de capítulo o sección.

Agradecemos a M. Buchner y W. Wilson su ayuda con el primer borrador del libro y a I. Fáry y R. Gulliver por los apéndices. También agradecemos a los estudiantes del curso de Math 104A-B en Berkeley, particularmente a E. Wong, J. Lim, J. Wing y J. Seitz, por haber atrapado a tiempo un buen número de pequeños errores y haber puntualizado algunas cuestiones de estilo. Damos las gracias a nuestros colegas de cuyos exámenes anteriores pudimos derivar muchos problemas. Varios colegas nuestros merecen una mención especial, particularmente P. Chernoff, I. Fáry, R. Gulliver y M. Mayer, por leer partes del manuscrito y sugerir diversas mejoras. Las secciones 5.9 y 10.4 y varios problemas fueron adaptados de las notas de clase de P. Chernoff. Los demás ayudantes, A. Erickson, A. Hausknecht, D. Heifetz y J. Macrae, colaboraron en partes del manuscrito, eliminaron muchos errores y revisaron y prepararon las respuestas de la mayoría de los problemas. También recibimos ayuda de M. McCracken, W. A. J. Luxemburg y R. Graff. Agradecemos a I. Workman su excelente trabajo de mecanografiado del manuscrito y a N. Lee su apoyo moral.

Por último, damos las gracias a R. Abraham, K. McAloon, A. Tromba y M. O'Nan, de Eagle Mathematics Incorporated (una organización de autores matemáticos) por su sugerencia de que escribiéramos este libro y su apoyo posterior.

Jerrold E. Marsden

Octubre de 1973

Introducción: conjuntos y funciones

El estudiante que desee usar este libro con buenos resultados debe tener sólidos conocimientos de cálculo elemental y álgebra lineal y un poco de experiencia en cálculo de varias variables. En general, se obtiene una preparación adecuada después de cursar dos años de matemáticas. También son necesarios conocimientos básicos de conjuntos y funciones, cuyos conceptos fundamentales se resumen en esta introducción. Conviene leer brevemente este material y consultarlo cuando se necesite.

La teoría de conjuntos es el punto de partida de gran parte de la matemática y constituye, por sí misma, un tema amplio y complicado. Con objeto de ser breves y lograr una mejor comprensión, comenzamos nuestro estudio de manera intuitiva. El lector que esté interesado en las sutilezas de la teoría de conjuntos puede consultar el suplemento que aparece al final de esta introducción.

Llamamos ***conjunto*** a una colección de "objetos" o "cosas" denominados ***elementos*** del conjunto. Por ejemplo, la colección de enteros positivos 1, 2, 3, . . . forma un conjunto. De la misma manera, los números racionales (las fracciones) p/q forman un conjunto. Si S es un conjunto y x es un elemento de S, escribimos $x \in S$. Decimos que un conjunto A es ***subconjunto*** de S si cada elemento de A es también un elemento de S; simbólicamente, esta relación se denota $(x \in A) \Rightarrow (x \in S)$, donde $\Rightarrow$ significa "implica". Cuando A es un subconjunto de S, escribimos $A \subset S$. En algunas ocasiones se usa la notación $A \subseteq S$ en lugar de $A \subset S$. También podemos definir la igualdad de conjuntos afirmando que $A = B$ significa $A \subset B$ y $B \subset A$; es decir, A y B tienen los mismos elementos. El ***conjunto vacío***, que simbolizaremos con $\varnothing$, es un conjunto sin elementos. Por ejemplo, el conjunto de los enteros n tales que $n^2 = -1$ es vacío.

Un método para describir un conjunto es enumerar sus elementos entre llaves. Así, escribimos $\mathbb{N} = \{1, 2, 3, \ldots\}$ para denotar el conjunto de todos los enteros positivos y $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ para el conjunto de todos los enteros. Un ejemplo de un subconjunto de $\mathbb{N}$ es el conjunto de todos los números pares; se escribe

$$A = \{2, 4, 6, \ldots\} = \{x \in \mathbb{N} \mid x \text{ es par}\} \subset \mathbb{N}.$$

Leemos $\{x \in \mathbb{N} \mid x \text{ es par}\}$ como "el conjunto de todos los elementos x de $\mathbb{N}$ tales que x es par". Debemos hacer una importante aclaración sobre la notación: si S es un con-

junto y $a \in S$, entonces $\{a\}$ denota el subconjunto de S que consta únicamente del elemento a; así, $\{a\} \subset S$, mientras que $a \in S$.

Sea S un conjunto dado y sean $A \subset S$ y $B \subset S$. Definimos $A \cup B = \{x \in S \mid x \in A$ o $x \in B\}$, lo cual se lee "el conjunto de todos los $x \in S$ que son elementos de A o de B (o de ambos)". El conjunto $A \cup B$ se denomina ***unión*** de A y B. De manera similar, podemos formar la unión de una familia de conjuntos. Por ejemplo, sean $A_1, A_2, \ldots$ subconjuntos de S y sea $\cup_{i=1}^{\infty} A_i = \{x \in S \mid x \in A_i$ para algún $i\}$; esto también se expresa $\cup \{A_1, A_2, A_3, \ldots\}$. Obsérvese que $A \cup B$ es el caso especial en que $A_1 = A$, $A_2 = B$ y $A_i = \varnothing$ para $i > 2$. Las intersecciones $A \cap B = \{x \in S \mid x \in A$ y $x \in B\}$ y $\cap_{i=1}^{\infty} A_i = \{x \in S \mid x \in A_i$ para todo $i\}$ se forman de manera similar. La figura I-1 ilustra estas operaciones.

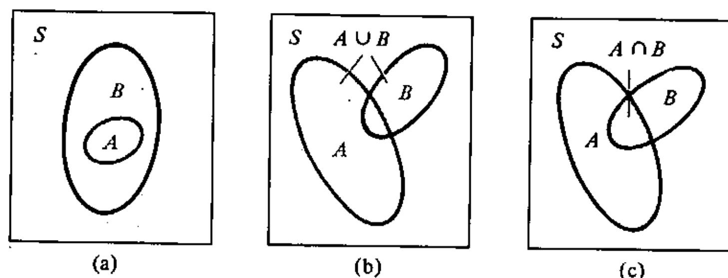


FIGURA I-1 (a) Subconjunto; (b) unión; (c) intersección

Dados $A \subset S$ y $B \subset S$, formamos el ***complemento*** de A relativo a B definiendo

$$B \backslash A = \{x \in B \mid x \notin A\},$$

donde $x \notin A$ significa que x no pertenece al conjunto A. Véase la figura I-2.

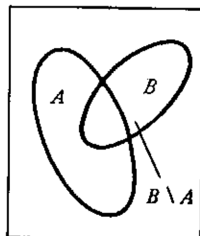


FIGURA I-2 Complemento

Como veremos en el ejemplo resuelto **I.1** al final de esta sección, $B\backslash(A_1 \cup A_2) = (B\backslash A_1) \cap (B\backslash A_2)$ y $B\backslash(A_1 \cap A_2) = (B\backslash A_1) \cup (B\backslash A_2)$ para cualesquiera $A_1, A_2, B \subset S$. Éste es un ejemplo de "identidad de conjuntos". En los ejercicios se dan otros ejemplos.

Dados los conjuntos A y B, definimos el ***producto cartesiano*** $A \times B$ de A y B como el conjunto de todos los ***pares ordenados*** (a, b) con $a \in A$ y $b \in B$, es decir, $A \times B = \{(a, b) \mid a \in A \text{ y } b \in B\}$. Véase la figura I-3.

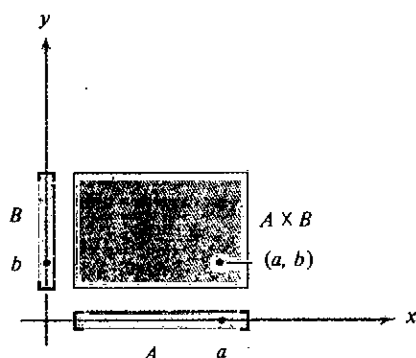


FIGURA I-3 Producto cartesiano

Sean S y T conjuntos dados. Una ***función*** $f: S \to T$ consta de dos conjuntos S y T junto con una "regla" que asigna a cada $x \in S$ un elemento específico de T, que denotaremos $f(x)$. A veces se escribe $x \mapsto f(x)$ para indicar que la imagen de x es el elemento $f(x)$. Por ejemplo, la función $f(x) = x^2$ se puede describir diciendo $x \mapsto x^2$. La figura I-4 representa esta función con $S = T$, el conjunto de todos los números reales; este conjunto se simboliza con $\mathbb{R}$ y se estudiará con más cuidado en el capítulo 1. Por ahora lo usaremos de una manera informal.

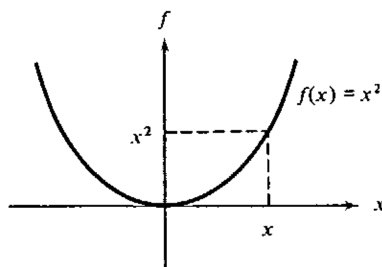


FIGURA I-4 La función $x \mapsto x^2$

Nota. En este libro, los términos "función", "aplicación" y "transformación" son sinónimos.

En una función $f: S \to T$, el conjunto S se llama ***dominio*** u ***origen*** de f y el conjunto T se llama ***destino*** de f. El ***rango*** o ***imagen*** de f es el subconjunto de T definido como $f(S) = \{f(x) \in T \mid x \in S\}$. La ***gráfica*** de f es el conjunto $\{(x, f(x)) \in S \times T \mid x \in S\}$, como en la figura I-5.

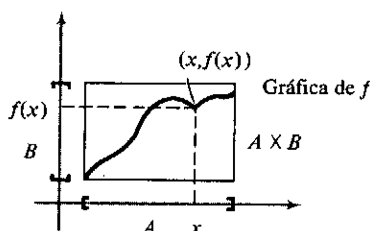


FIGURA I-5 Gráfica de una función

Es posible que quienes sean muy estrictos con la fundamentación lógica se opongan al uso de palabras del lenguaje cotidiano, como "regla", y prefieran definir una función de S a T como un subconjunto R de $S \times T$ con las siguientes propiedades:

1. Todo elemento de S debe aparecer como la primera componente de algún elemento de R, y

2. Dos elementos de R con la misma primera componente son idénticos; es decir, la primera componente x determina la segunda componente $f(x)$, como en la figura I-5.

Decimos que una función $f: S \to T$ es ***inyectiva*** o ***unívoca*** si $x_1 \neq x_2$ implica $f(x_1) \neq f(x_2)$. Por tanto, una función es inyectiva cuando no se dé el caso de que dos elementos distintos de S tengan como imagen el mismo elemento de T. Lo anterior equivale a decir que *f es inyectiva cuando para cada $y \in T$, la ecuación $f(x) = y$ tiene como mucho una solución $x \in S$.* Un ejemplo extremo de una función que no es inyectiva (si S tiene más de un elemento) es una función constante, es decir, una función $f: S \to T$ tal que $f(x_1) = f(x_2)$ para todo $x_1, x_2 \in S$. Véase la figura I-6.

Nota. En las definiciones se acostumbra escribir "si" en lugar de "si y sólo si"; esto último suele escribirse "sii", o $\Leftrightarrow$. En los teoremas es absolutamente necesario distinguir entre "si", "sólo si" y "sii".

Decimos que $f: S \to T$ es ***suprayectiva*** o que es una función de S ***sobre*** T cuando para todo elemento $y \in T$ existe un elemento $x \in S$ tal que $f(x) = y$; en otras palabras,

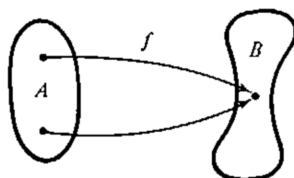


FIGURA I-6 Función constante

cuando el rango es igual al destino. Otra manera de decir esto es que *para cada* $y \in T$, *la ecuación* $f(x) = y$ *tiene al menos una solución* $x \in S$. Debemos hacer notar que la elección de S y T es parte de la definición de f, y que el hecho de que f sea inyectiva o supráyectiva depende de dicha elección. Por ejemplo, sea f definida por $f(x) = x^2$; entonces f es inyectiva y suprayectiva cuando S y T son el conjunto que consta de los números reales tales que $x \geq 0$; es inyectiva pero *no* suprayectiva cuando S son todos los x tales que $x \geq 0$ y T es el conjunto de todos los números reales, y *no* es inyectiva *ni suprayectiva* cuando S y T son ambos el conjunto de los números reales.

Para $f : S \to T$ y $A \subset S$, definimos $f(A) = \{ f(x) \in T \mid x \in A \}$ y para $B \subset T$ definimos $f^{-1}(B)$ como el conjunto $\{x \in S \mid f(x) \in B\}$. Llamamos a $f(A)$ la ***imagen*** de A bajo f y a $f^{-1}(B)$ la ***imagen inversa*** o ***preimagen*** de B bajo f.

Nota. Podemos construir $f^{-1}(B)$ para cualquier conjunto $B \subset T$ aun cuando f no sea inyectiva.

Si $f: S \to T$ es inyectiva y suprayectiva, entonces *para cada* $y \in T$ existe una solución única $x \in S$ de la ecuación $f(x) = y$. Por lo tanto, existe una función única, a la cual denotaremos con $f^{-1} : T \to S$ (que no se debe confundir con la operación $f^{-1}(B)$ definida en el párrafo anterior, o con $1/f$), tal que $f(f^{-1}(y)) = y$ para todo $y \in T$ y $f^{-1}(f(x)) = x$ para todo $x \in S$. Llamamos a f^{-1} la ***función inversa*** de f. Una función que sea inyectiva y suprayectiva se denomina también ***biyección*** o ***correspondencia biunívoca***.

Nota. En cálculo aprendemos lo importante que es elegir bien el dominio (origen) para construir la inversa de una función. Por ejemplo, para construir $\text{sen}^{-1} \equiv \text{arcsen}$, reducimos el dominio y consideramos sen como la función $\text{sen}:[-\pi/2, \pi/2] \to [-1, 1]$, que es una biyección. Entonces $\text{sen}^{-1}: [-1, 1] \to [-\pi/2, \pi/2]$ está definida. En cualquier libro de cálculo se pueden encontrar más ejemplos.

La función $f: S \to S$ definida por $f(x) = x$ para todo $x \in S$ se llama ***función identidad*** sobre S. Debemos distinguir la función identidad para conjuntos diferentes. Por ejemplo, algunas veces usamos la notacion I_s para la función identidad sobre S. Evidentemente, I_s es inyectiva y suprayectiva.

Para dos funciones $f : S \to T$ y $g : T \to U$ se define la ***composición*** $g \circ f : S \to U$ mediante $(g \circ f)(x) = g(f(x))$, como se muestra en la figura I-7. Por ejemplo, si $f : \mathbb{R} \to \mathbb{R}$ está definida por $f(x) = x^2$ y $g : \mathbb{R} \to \mathbb{R}$ por $g(x) = x + 3$, entonces $g \circ f : x \mapsto x^2 + 3$ y $f \circ g : x \mapsto (x + 3)^2$ (en este caso, S, T y U son el conjunto de todos los números reales). En particular, obsérvese que $f \circ g \neq g \circ f$.

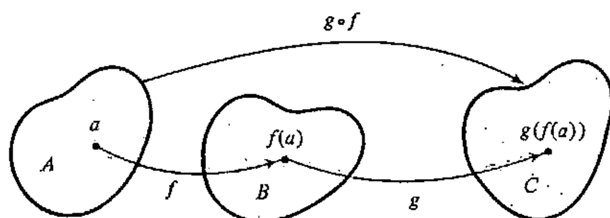


FIGURA I-7 Composición de funciones

> **Nota.** En cálculo aprendemos que la operación de composición es importante, entre otras cosas, para la regla de la cadena. La misma importancia tiene en este libro.

Algunas veces queremos concentrar nuestra atención sólo en algunos de los elementos sobre los que está definida una función. Este proceso se llama ***restricción*** de una función. De manera más formal, si tenemos una transformación $f : S \to T$ y $A \subset S$, consideramos una nueva función, que denotamos $f|A : A \to T$, definida por $(f|A)(x) = f(x)$ para todo $x \in A$. Llamamos a $f|A$ la ***restricción*** de f a A y decimos que f es una ***extensión*** de $f|A$.

Decimos que un conjunto A es ***finito*** si podemos presentar explícitamente todos sus elementos de la siguiente manera: $A = \{a_1, a_2, \ldots, a_n\}$ para algún entero n. Un conjunto que no es finito se llama ***infinito***. Por ejemplo, el conjunto de los enteros positivos $\mathrm{N} = \{1, 2, \ldots\}$ es un conjunto infinito. Puede ser difícil determinar si un conjunto infinito tiene más elementos que otro conjunto infinito. Por ejemplo, no está claro a primera vista si hay más números racionales o irracionales. Para precisar esta noción, decimos que dos conjuntos A y B tienen el ***mismo número de elementos*** (o tienen la misma ***cardinalidad***) si existe una función $f : A \to B$ que es inyectiva *y* suprayectiva. Si un conjunto es finito, o bien si es infinito pero tiene el mismo número de elementos que el conjunto de los enteros positivos $\{1, 2, \ldots\}$, se dice que es ***numerable***. Del resto de los conjuntos diremos que son infinitos no numerables, o simplemente ***no numerables***. Un ejemplo de un conjunto no numerable es el conjunto de todos los números reales entre 0 y 1 (lo demostraremos en el capítulo 1).

Sea S un conjunto. Una ***sucesión*** en S es una función $f : \mathrm{N} \to S$, que asocia a cada entero n un elemento de S, a saber $f(n)$. A menudo se suele hacer a un lado el hecho de

que tenemos una función, considerando una sucesión como el conjunto de elementos de la imagen, digamos $x_1, x_2, x_3, \ldots$, o, alternativamente, escribimos "la sucesión x_n" o bien $(x_n)_{n=1}^{\infty}$. La sucesión $y_1, y_2, \ldots$ es, a su vez, una ***subsucesión*** de $x_1, x_2, \ldots$, si existe una función $g : \mathbf{N} \to \mathbf{N}$ tal que para todo $i \in \mathbf{N}$, $y_i = x_{g(i)}$, y para $i < j$, $g(i) < g(j)$. En otras palabras, obtenemos una sucesión "desechando" elementos de la sucesión original y ordenando de la manera natural los elementos que quedan. Por ejemplo, la sucesión $y_n = (2n)^2$ es una subsucesión de $x_n = n^2$; en este caso, $g(n) = 2n$. Se acostumbra a expresar una subsucesión x_n en la forma x_{n_i}, donde la notación $g(i) = n_i$ nos recuerda que los índices n_i se eligen de entre los n.

Un método importante para demostrar afirmaciones que dependen de un índice que recorre los enteros positivos $\mathbf{N}$ es la técnica llamada ***inducción***. Una propiedad $P(i)$ es verdadera para todo $i \in \mathbf{N}$ si:

1. $P(1)$ es verdadera (caso base), y

2. Para todo $n \in \mathbf{N}$, si $P(n)$ es verdadera entonces $P(n + 1)$ es verdadera (paso de inducción).

La misma técnica se puede aplicar a $\{0, 1, \ldots\}$ si sustituimos el caso base por:

1′. $P(0)$ es verdadera.

Tendremos más que decir sobre las bases para el conjunto de los números naturales en **§1.1**.

Suplemento sobre los axiomas de la teoría de conjuntos[1]

No existe campo de la matemática formal contemporánea que no utilice conceptos de la teoría de conjuntos. Por este motivo hemos comenzado el libro con la teoría de conjuntos. El propósito de este suplemento es ayudar a llenar el hueco que existe entre este texto y los cursos de teoría de conjuntos más formales que usan libros como *Naive Set Theory*, de Halmos.[2] Una introducción a la teoría de conjuntos debe tener en cuenta los siguientes puntos:

1. El concepto de conjunto es tan básico que es imposible de definir en términos de nociones más básicas.

2. Por esta razón, especificamos el concepto de conjunto con axiomas, pero el método axiomático puede no resultarle familiar al estudiante.

3. La teoría axiomática de conjuntos utiliza la lógica, pero es posible que tampoco algunos conceptos de lógica le sean familiares.

En virtud de lo anterior, la manera más eficaz de abordar el tema, y la que usaremos en este libro, es comenzar trabajando con el concepto intuitivo de conjunto y regresar a

[1] Este suplemento se escribió con la ayuda de István Fáry.

[2] Halmos, Paul R., 1960, *Naive Set Theory*. Nueva York: D. Van Nostrand Co.

los fundamentos después. Cuando se usa este método surge la pregunta de si se debe estudiar lógica primero o estudiar la teoría axiomática de conjuntos sin la lógica formal. Nosotros hemos elegido la segunda opción.

Este plan se corresponde con el desarrollo histórico: primero apareció la teoría de conjuntos basada en conceptos intuitivos; a continuación, la crítica a este procedimiento inspiró la fundamentación axiomática, y, por último, un intenso debate sobre este método impulsó nuevos avances en lógica. Tal vez sea útil, por lo tanto, dedicar unas líneas a la historia de este tema.

Sobre la historia de la teoría de conjuntos

La teoría de conjuntos es una de las áreas más básicas de la matemática. Contiene proposiciones sobre conjuntos finitos, pero la importancia de la teoría de conjuntos radica en que puede manejar conjuntos infinitos y en que se puede desarrollar de manera sistemática. En este sentido, el fundador de la teoría de conjuntos fue Georg Cantor (1845–1918), quien publicó sus importantes artículos justo antes del inicio del siglo veinte. Su trabajo suscitó un acalorado debate, y matemáticos de renombre discreparon en algunos puntos fundamentales.

Cantor fue el primero en descubrir propiedades de los conjuntos infinitos en relación con su trabajo sobre series trigonométricas (véase el capítulo 10). Una serie trigonométrica es una suma de la forma

$$\sum_{k=0}^{\infty}(a_k \cos kx + b_k \operatorname{sen} kx).$$

Las propiedades de convergencia de estas series son una cuestión delicada, y definir conjuntos de puntos de acuerdo con el comportamiento de la serie conduce a tipos muy generales de conjuntos de números. Por este motivo, Cantor estudió primero los conjuntos de números reales, pero pronto descubrió que tenía que trabajar con conjuntos infinitos en general.

En uno de sus artículos proporciona la siguiente "definición" del concepto de conjunto:

> *Entendemos por "conjunto" cualquier agrupación M de objetos m (los cuales se llamarán "elementos" de M) bien definidos y distinguibles de nuestra intuición o de nuestras ideas en un todo.*[3] (1)

En la actualidad es costumbre "avergonzarse" de la definición original de Cantor y decir que no es una definición. Aun así, existen muchas otras de las llamadas definiciones en otros campos que no tienen la claridad y precisión de (1). No obstante, dado que el concepto de conjunto es tan importante, no aceptaremos la definición de Cantor

---

[3] El texto original en alemán (*Collected Papers*, pág. 282) es: "Unter einer 'Menge' verstehen wir jede Zusammenfassung *M* von bestimmten wohlunterschiedenen Objekten *m* unserer Anschauung oder unseres Denkens (welche die 'Elemente' von *M* genannt werden) zu einem Ganzen."

como la última palabra, aunque por el momento usaremos (1) para esclarecer más nuestras ideas sobre conjuntos.

La primera idea es que "agrupamos" objetos y no consideramos el *orden* en que se toman. Por ejemplo, si hablamos acerca del "conjunto de los números naturales", no suponemos que los elementos de este conjunto están dados en algún "orden", aun cuando exista un "orden natural" para los enteros. Por motivos prácticos, podemos dar una lista de los elementos en algún orden, pero esto no tiene nada que ver con el conjunto mismo.

Las palabras "objetos bien definidos y distinguibles" en (1) señalan otro aspecto del concepto de conjunto: los *elementos* del conjunto no "aparecen dos veces"; por ejemplo, un conjunto que consta de 2, 2, 2, 3 contiene el 2 y el 3 y nada más. Por lo tanto, un conjunto "contiene" objetos que le "pertenecen"; hay otros objetos que pueden no pertenecer al conjunto. Por ejemplo, 1003 pertenece al conjunto de los números naturales (enteros positivos), mientras que 3.14159 no pertenece a él.

Por último, la palabra "todo" que aparece al final de la definición de (1) se refiere al hecho de que los conjuntos mismos se tratan como objetos; por ejemplo, pueden ser elementos de otros conjuntos. Así pues, podemos considerar conjuntos cuyos elementos son conjuntos. De hecho, los de este tipo están entre los más importantes de la teoría de conjuntos.

Hagamos ahora una crítica de la definición de Cantor. Considérese la definición siguiente: un número p es un ***número primo*** si $p \neq 1$ y $\pm 1, \pm p$ son los únicos divisores de p. En esta definición, el concepto de "número primo" está definido en términos de otros conceptos (enteros, divisor, $+1$, -1, $-p$), y suponemos que éstos se conocen o se definieron sin usar el concepto de número primo. La definición, pues, reduce el concepto de números primos a los otros conceptos. La definición también nos dice lo que debemos hacer para determinar si 1003 es o no un número primo (no lo es, porque es divisible por 17). Veamos si (1) satisface dicho criterio. Tenemos en este enunciado una variedad de otros conceptos: "agrupación", "bien definidos", "distinguibles", "todo" (por no mencionar nuestra "intuición", "nuestras ideas"). Es lícito preguntar qué concepto es más sencillo: "conjunto" o "agrupación". (De hecho, el vocablo alemán "Zusammenfassung" suena mejor, pero no escapa a la crítica.) De manera similar, podemos cuestionar cada uno de los demás conceptos y preguntar si son más sencillos que el concepto de conjunto. La definición de Cantor también recibió críticas sobre la base de que no excluye conjuntos contradictorios, como veremos en la sección siguiente. Motivado por estas críticas, Cantor descubrió el germen de la descripción axiomática en un artículo posterior.

Lógica

No estudiaremos la lógica desde un punto de vista crítico y sofisticado. Probablemente es justo decir que la base de nuestro pensamiento racional es la siguiente creencia: Si partimos de premisas verdaderas y hacemos deducciones correctas a partir de ellas, llegaremos a conclusiones verdaderas. Podemos rechazar este enunciado, pero no lle-

garíamos muy lejos en matemáticas. Si tomamos en serio esta creencia (como lo hacemos en matemáticas), entonces podemos alcanzar resultados bastante sofisticados. Por ejemplo, supongamos que 2, 3, 5, 7, 11, 13 y 17 son los únicos números primos ≥ 2. A continuación formemos el número $n = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 + 1$. Es fácil ver que n no es divisible por ningún primo ≤ 17 (si n se factorizase como un primo por un entero obtendríamos 1 como un producto no trivial de enteros, lo cual es imposible). Por lo tanto, n es primo o bien tiene un divisor que es un primo ≥ 18. Como la conclusión contradice por completo la premisa, no es posible que ambas sean ciertas, y *como nuestro razonamiento ha sido correcto*, la premisa debe ser falsa: existe un número primo ≥ 18. No resulta sorprendente, pues 19 resulta ser ese primo, pero hemos llegado a la conclusión razonando y no por experiencia. Más importante aún es el hecho de que el mismo tipo de argumentación demuestra que existe un número primo mayor que cualquier número primo dado; *es decir, existe una cantidad infinita de primos*. Este razonamiento, llamado *reducción al absurdo*, se usa con frecuencia.

Existen oraciones en nuestro idioma en las cuales podemos suprimir una palabra, escribir x en su lugar y seguir teniendo un enunciado con sentido. Por ejemplo, en la frase "dos es menor que cinco", si borramos "dos" y escribimos "x" obtenemos el enunciado "x es menor que cinco". Dicha combinación de palabras se llama *función proposicional* o *condición* y se denota por $S(x)$. Si escribimos "siete" en lugar de x obtenemos un enunciado falso, y si escribimos "tres" en lugar de x, el enunciado es verdadero; así que $S(x)$ tiene sentido si x es un entero y es verdadero para algunos enteros y falso para otros. Dada una condición arbitraria $S(x)$, podemos

Tomar todos los objetos cuyo nombre, sustituido en lugar de x en $S(x)$, nos dé un enunciado verdadero. (2)

Se entiende que x puede aparecer varias veces y que la sustitución se hace de manera consistente (así x sólo "reserva el lugar"). A partir de la definición (1) obtenemos, por tanto, un conjunto. Existe una notación usual para este conjunto:

$$\{x \mid S(x)\}. \tag{3}$$

A pesar de que (2) es consistente con (1) y con el concepto cotidiano de conjunto, caemos en contradicciones si usamos (2) sin criterio. Tomemos el ejemplo siguiente:

El conjunto que no se contiene a sí mismo como elemento. (4)

La frase parece ser correcta; después de todo, ¿quién ha visto un conjunto que se contenga a sí mismo como elemento? Borremos "el conjunto que", y escribamos "x" para obtener el enunciado equivalente:

$$S(x) = x \text{ no se contiene a sí mismo como elemento.} \tag{5}$$

Luego tomemos el conjunto correspondiente tal como se describe en (2). Llamémoslo M, el nombre que Cantor le dio (M de "Menge"). Consideremos la pregunta: ¿contiene

M a M? Si no lo contiene, lo debe contener, por el enunciado que lo define; si lo contiene, entonces no lo debe contener, en virtud del mismo enunciado.

Esta propiedad de la construcción (2), señalada por primera vez por Bertrand Russell, es traumática y desalentadora. Cuando analizábamos la definición de Cantor, sugeríamos que no era en realidad tan mala y que ciertamente ayuda a aclarar las ideas. Ahora vemos que, al mismo tiempo, permite construir el conjunto imposible M.

El ejemplo del conjunto M indica que existe algún aspecto incorrecto inherente al concepto mismo de conjunto, o al menos a la parte relativa a conjuntos "grandes". De hecho, M es de los conjuntos más grandes que hay: contiene *todos* los conjuntos "decentes". Sin embargo, el tipo de contradicción que surge en relación con (5) se conoce muy bien en la lógica clásica. Mencionemos primero un ejemplo que se puede formular en téminos de conjuntos "pequeños". Sea N el conjunto de los hombres que viven en una pequeña aldea. Supongamos que el barbero de la aldea enuncia: "Yo afeitaré a $x \in N$ si y sólo si x no se afeita a sí mismo." Parece ser que este enunciado define un subconjunto $P \subset N$. Sin embargo, la pregunta de si el barbero pertenece a P conduce a la llamada "paradoja del barbero": "me afeitaré a mí mismo si no me afeito a mí mismo".

Nota. En un lenguaje más moderno, considérese la "lámpara de Russell" en el cuadro de mandos de un automóvil. Ésta es una luz que se enciende cuando cualquiera de las luces del cuadro de mandos se funde. ¿Qué ocurre si se funde la lámpara de Russell?

Los lógicos griegos estudiaron de manera profunda este dilema, pero ellos no usaron el concepto de conjunto. Por lo tanto, la contradicción parece ser independiente de la noción de conjunto. Esto se confirma, al parecer, con la siguiente paradoja.

Supongamos que en una de mis clases un estudiante dice:

La última frase escrita en la pizarra es falsa. (6)

Por desgracia, esto puede ocurrir. De ser así, en general hago lo siguiente: leo otra vez la frase; si el estudiante tiene razón, pido disculpas, la borro y la corrijo; si, por el contrario, el estudiante se ha equivocado la leo en voz alta y la dejo tal como está en la pizarra. Para ser más específicos, supongamos ahora que mi clase trata sobre teoría de conjuntos y que he escrito las oraciones de la (1) a la (6); sólo estas oraciones y tal como están en la pizarra, en el mismo orden. Si un estudiante ahora dice: "La última frase de la pizarra es falsa", no sé qué debo hacer. Si el estudiante tiene razón, entonces (6) es falsa, lo cual significa que está correcta; por tanto, el estudiante está equivocado, pero en ese caso el enunciado o frase está correcto, lo cual significa que es falso.

Sería interesante profundizar en estos aspectos de la lógica, pero nuestro propósito sólo era indicar por qué es recomendable restringir la forma de los enunciados cuando se definen subconjuntos de un conjunto en nuestra teoría axiomática de conjuntos.

El lenguaje de los axiomas

En una exposición completa y avanzada de los axiomas de la teoría de conjuntos es necesario utilizar la lógica formal. Así, al menos una parte del lenguaje de la teoría de conjuntos está formalizada. Vamos a describir ahora esta parte del lenguaje sin llevar a cabo una formalización.

Existen dos tipos básicos de enunciados, a saber, los de pertenencia,

$$x \in A, \tag{7}$$

y los de igualdad

$$A = B. \tag{7'}$$

Todos los demás se obtienen a partir de este tipo de enunciados ***atómicos***, mediante la aplicación reiterada de los operadores usuales de la lógica, sujetos a las reglas de la gramática y ajenos a la ambigüedad.

Para hacer la definición explícita es necesario agregar una lista de los operadores "usuales" de la lógica y de las reglas de sintaxis. Nuestra lista de operadores lógicos será

$$\begin{cases} \text{no} \\ \text{y} \\ \text{o (en su sentido no exclusivo)} \\ \text{si—entonces— (que significa implica)} \\ \text{si y sólo si (que se abrevia sii)} \end{cases} \tag{8}$$
$$\begin{cases} \text{para algún— (existe)} \\ \text{para todo—} \end{cases}$$

Obsérvese que "no" actúa sobre un solo enunciado, mientras que los otros cuatro operadores actúan sobre dos enunciados (S y $T, \ldots, S$ sii T) y los últimos dos actúan sobre predicados (para algún x, $S(x)$ es cierta, etcétera). Esta lista es redundante: puede demostrarse en lógica que los cinco primeros pueden ser reemplazados por menos operadores. [Por ejemplo, podemos borrar "y" de la lista; en lugar del enunciado "S y T" donde S y T son enunciados, podemos decir "no (no S o no T)". Esto resulta confuso en el lenguaje cotidiano, pero resulta muy simple con el simbolismo lógico adecuado. Sin embargo, como no queremos usar lógica formal, usaremos la lista más larga (8).] En nuestra lista, los primeros cinco operadores se llaman ***conectivos lógicos***, y los dos últimos se llaman ***cuantificadores***. En el formalismo usual, "para alguna x" se escribe $\exists x$ y "para toda x" se denota $\forall x$. La conexión entre estos dos cuantificadores es la siguiente: la negación de "para alguna x, $S(x)$ se cumple" es "para toda x, no $S(x)$ se cumple"; la negación de "para toda x, $S(x)$ se cumple" es "existe una x para la cual no $S(x)$ se cumple". Ésta es posiblemente la idea *más importante* que se debe aprender en esta parte. A menudo la conexión entre los dos cuantificadores aparece de la siguiente forma. Queremos demostrar un enunciado:

$$\text{Para toda } \varepsilon_0 > 0, \cdots \text{ es cierta.} \tag{9}$$

Su negación es

$$\text{Existe } \varepsilon_0 > 0, \text{ tal que} \cdots \text{ no es cierta.} \qquad (10)$$

Si podemos deducir una contradicción de (10) habremos demostrado (9). En lo que a las reglas de construcción de enunciados se refiere, usaremos las siguientes convenciones:

1. Al usar "no", éste se pone antes del enunciado y se encierra todo entre paréntesis (el motivo de los paréntesis es garantizar la falta de ambigüedades).
2. Cuando se usen "y", "o" o "si, y sólo si" se ponen entre las dos oraciones o enunciados a los que se aplica y se encierra todo entre paréntesis.
3. Se sustituyen los guiones en "si—entonces—" por oraciones o enunciados y se encierra el resultado entre paréntesis.
4. Se sustituyen los guiones en "para algún—" o en "para todo—" por una variable, al resultado le sigue una oración y se encierra todo entre paréntesis. [No es incorrecto que la variable utilizada no aparezca en el enunciado. Según la convención habitual, "para algún y ($x \in A$)" significa "($x \in A$)". De la misma manera, tampoco es incorrecto que la variable haya aparecido ya en "para algún—" o "para todo—". "Para algún x ($x \in A$)" significa lo mismo que "Para algún y ($y \in A$)"; un cambio de notación juicioso siempre evitará colisiones alfabéticas.]

Los axiomas

En lugar de dar una definición del concepto de "conjunto A" y del de "pertenencia a un conjunto", denotado $a \in A$, daremos propiedades de dichos conceptos. La enumeración de propiedades es la principal característica del método axiomático. Pasemos, pues, a enunciar los axiomas, junto con algunas observaciones.

1. Axioma de extensión *Dos conjuntos son iguales si y sólo si tienen los mismos elementos.*

Este axioma significa, en particular, que si queremos probar que $A = B$, entonces debemos probar que $x \in A$ implica $x \in B$ y que $x \in B$ implica $x \in A$. Este punto es tan importante que vale la pena tener una notación para el caso en que la mitad del enunciado —digamos, la primera mitad— sea válida. En ese caso escribiremos $A \subset B$. Ésta será una relación entre conjuntos; no es un concepto indefinido, sino que se *define* en términos de "conjunto" y "pertenencia". Véase el ejemplo **I.1** al final de esta introducción para una aplicación concreta de este axioma.

2. Axioma de especificación *Para cada conjunto A y para cada condición $S(x)$ existe un conjunto B cuyos elementos son exactamente aquellos elementos x de A para los cuales $S(x)$ es verdadera.*

Introducimos la notación

$$\{x \in A \mid S(x)\} \tag{11}$$

para denotar al conjunto B. Obsérvese que (11) es igual a nuestro conjunto (3) salvo por el hecho de que ahora no formamos el conjunto de todos los objetos que satisfacen cierta condición sino sólo de aquellos que ya son elementos de algún conjunto (el conjunto A en (11)). Esto nos permite, por ejemplo, formar conjuntos de números reales de manera totalmente arbitraria, tales como

$$I = \{x \in \mathbb{R} \mid a \leq x \leq b\}, \tag{12}$$

donde nuestra condición $S(x)$ es $a \leq x \leq b$, siempre que sepamos que $\mathbb{R}$ es un conjunto (lo que todavía no es consecuencia de los axiomas 1 y 2). El conjunto (11) más simple se puede formar con *el enunciado atómico* (7); así, obtenemos $A = \{x \in A \mid x \in A\}$, de donde A es un subconjunto de A. Si *no* hay elementos de A que satisfagan la condición $S(x)$, entonces (11) describe el ***conjunto vacío*** $\varnothing$. Siempre podemos plantear una condición imposible, por ejemplo, $x \notin$ A. Entonces $\varnothing = \{x \in A \mid x \notin A\}$. Por lo tanto, si existe al menos un conjunto, entonces existe también un conjunto que no tiene elementos (nuestros axiomas no dicen aún que existan conjuntos; postularemos tal enunciado después).

Sobre la base del axioma 2, introducimos la importante operación teórica de intersección de conjuntos. Dados dos conjuntos A y B, escribimos $\{x \in A \mid x \in B\}$; esta operación se denota $A \cap B$, como es bien conocido. El conjunto $B \cap A$ es $\{x \in B \mid x \in A\}$, el cual es claramente igual a $A \cap B$. La operación más general es la intersección de una colección de conjuntos C (en lugar de un *conjunto* de conjuntos se suele decir *colección* de conjuntos, pero, para nosotros, "colección" será sinónimo de "conjunto"): supongamos que C es un conjunto, y, si $A \in C$, entonces A también es un conjunto. Definimos:

$$\bigcap\{A \mid A \in C\} = \{x \in A_0 \mid A_0 \in C \ \text{ y } \ x \in A \text{ para todo } A \in C\}. \tag{13}$$

Por lo tanto, x es un elemento de la intersección si pertenece a todos los conjuntos que pertenecen a C. $A \cap B$ corresponde al caso en que C contiene dos elementos, uno igual a A y el otro igual a B. Podemos etiquetar los elementos de C con números enteros, de tal manera que $C = \{A_n\}$, y, en ese caso, escribiremos

$$\bigcap_{n=1}^{\infty} A_n = \{x \in A_1 \mid x \in A_n \text{ para todo } n\}. \tag{14}$$

Está claro que (14) es el conjunto (13) para este caso particular (en algunos casos los elementos de C no se pueden etiquetar de esta manera).

3. Axioma de apareamiento *Para cualquier par de conjuntos, existe un conjunto que los contiene a ambos.*

4. Axioma de la unión *Para cualquier colección de conjuntos, existe un conjunto que contiene todos los elementos que pertenecen al menos a uno de los conjuntos de la colección dada.*

Si C es como en (13), escribimos

$$\bigcup\{A \mid A \in C\} \tag{15}$$

para denotar al conjunto postulado en el axioma 4. Las notaciones $A \cup B$ y $\bigcap_{n=1}^{\infty} A_n$ se usan en casos especiales, como ocurre con las intersecciones. Si queremos demostrar que $x \in \bigcap A_n$, debemos probar que $x \in A_n$ *para al menos un n*; si queremos demostrar que $x \in \bigcap A_n$, entonces debemos probar que $x \in A_n$ *para todo n*. Por lo tanto, los cuantificadores $\exists x$ y x están íntimamente relacionados con las operaciones $\cup$ y $\cap$ de la teoría de conjuntos.

Debemos ser cuidadosos y distinguir entre apareamiento y unión. El conjunto $\{A, B\}$ postulado por el axioma 3, tiene dos elementos, A y B, si $A \neq B$, y un solo elemento A si $B = A$ (este caso no se excluye). Por ejemplo, dado el conjunto $\varnothing$, podemos formar el conjunto $\{\varnothing, \varnothing\} = \{\varnothing\}$, el cual es un conjunto no vacío: tiene un elemento. El axioma 4 postula la existencia de $A \cup B$. En general, este conjunto no contiene ni A ni B como elementos. Sus elementos son elementos de A o de B. Por ejemplo, $\varnothing \cup \varnothing = \varnothing$ no tiene elementos; por lo tanto, es diferente de $\{\varnothing\}$.

Los axiomas 3 y 4 también *implican la existencia de un conjunto* $\{A, B, C\}$ *con tres elementos*. La manera de ver esto es formar $\{A, B\}$ y $\{C, C\} = \{C\}$ y luego el par $\{\{A, B\}, \{C\}\} = D$. Finalmente, tomamos la unión de los elementos de D. De manera similar, dados n conjuntos $A_1, \ldots, A_n$, podemos formar el conjunto $\{A_1, \ldots, A_n\}$.

5. Axioma de la potencia *Para cada conjunto existe una colección de conjuntos que contiene entre sus elementos todos los subconjuntos del conjunto dado.*

El axioma de la potencia es una herramienta básica de la teoría de conjuntos. Ya sabemos lo que son los conjuntos numerables. En el ejemplo **I.6** que aparece al final de esta sección mostraremos que si A es numerable, entonces el conjunto potencia $\mathcal{P}(A)$ no es numerable, y, de manera más general, que no hay biyecciones de A a $\mathcal{P}(A)$. Esto lo descubrió Cantor; la teoría de conjuntos, tal como la conocemos hoy, se inició con este descubrimiento. Si A es numerable, entonces existe una biyección de $\mathcal{P}(A)$ a $\mathbb{R}$, es decir, el conjunto de los números reales. Por lo tanto, si aceptamos la existencia de los enteros como un conjunto, los axiomas del 1 al 5 implican la existencia del conjunto $\mathbb{R}$ de los números reales, como veremos en **§1.2**. Pero los axiomas dados hasta ahora no postu-

lan la existencia de ningún conjunto, y mucho menos la existencia de conjuntos infinitos. Por ejemplo, el axioma 4 se entiende de esta manera: si tenemos una colección de conjuntos, entonces podemos formar la unión, ¡pero nadie ha dicho que tengamos ningún conjunto!

Antes de formular el último grupo de axiomas, estudiaremos más de cerca el problema de la existencia de conjuntos. Si entendemos los conjuntos en el sentido de la definición de Cantor (1), todos nuestros axiomas se cumplen. De ellos podemos deducir que algunos conjuntos formados a partir de la definición de Cantor no son conjuntos en el sentido de los axiomas; de manera más específica, dado un conjunto A podemos formar $B = \{x \in A \mid x \notin x\}$. Supongamos ahora que $B \in B$; entonces $B \notin B$, por definición; de donde se sigue que B no puede ser un elemento de sí mismo. En conclusión, $B \notin B$, y, en particular, $B \notin A$. Resumiendo: dado un conjunto A, podemos construir un conjunto B que *no* es elemento de A. Por lo tanto, los axiomas *impiden* la existencia de un conjunto A que contenga *todos* los conjuntos. Por el contrario, la definición de Cantor, admitiría un conjunto como ése. De manera similar, la contradicción referente al conjunto M de (5) demuestra que M no es un conjunto. El sistema axiomático cumple, entonces, nuestros propósitos. *Sobre la base de los axiomas podemos desarrollar parte de la teoría de conjuntos de Cantor, indispensable en matemáticas, y, al mismo tiempo, podemos evitar las contradicciones de la teoría de Cantor.*

Si sustituimos la palabra "conjunto" en los axiomas del 1 al 5 por la palabra "conjunto finito", tenemos enunciados consistentes. Como queremos introducir el concepto de "conjunto" con estos axiomas, debemos aceptar cualquier interpretación consistente con ellos. Por lo tanto, se necesita un axioma de infinitud.

Definición *Si x en un conjunto, entonces definimos $x^+ = x \cup \{x\}$ y lo llamamos el **sucesor** de x.*

6. Axioma de infinitud *Existe un conjunto que contiene el conjunto vacío ∅ y el sucesor de todos y cada uno de sus elementos.*

Éste es el tipo de axioma necesario para introducir los enteros. El siguiente axioma, que ha sido objeto de controversia en la historia de la teoría de conjuntos, asegura que de cualquier colección de conjuntos podemos "elegir" un representante de cada conjunto de la colección. Esto se enuncia de manera más precisa en el siguiente axioma.

7. Axioma de elección *Si A es una colección de conjuntos disjuntos dos a dos y no vacíos, entonces existe un conjunto C de elección tal que, para cualquier conjunto x de A, $x \cap C$ contiene un único elemento.*

Ya hemos estudiado el concepto de *par ordenado* (a, b) de los elementos a y b. Un par ordenado contiene un primer elemento (coordenada) a y un segundo elemento

(coordenada) b. El concepto de par ordenado se puede reducir al concepto de conjunto definiendo $(a, b) = \{\{a\}, \{a, b\}\}$, pero no entraremos aquí en detalles.

Si A y B son conjuntos dados, podemos formar el conjunto de *todos* los pares ordenados (a, b), donde $a \in A$ y $b \in B$;[4] este conjunto se denota $A \times B$. Por definición, una función $f: A \to B$ es un subconjunto de $A \times B$ tal que:

1. Dado $a \in A$ existe $b \in B$ tal que $(a, b) \in f$;
2. Si $(a, b_1) \in f$ y $(a, b_2) \in f$, entonces $b_2 = b_1$.

En lugar de $(a, b) \in f$, escribiremos $b = f(a)$. A partir de aquí podemos proceder a definir los siguientes términos, notaciones y conceptos relacionados con funciones: inyección, suprayección, biyección, restricción, extensión, $f(X)$ si $X \subset A$, $f^{-1}(Y)$ si $Y \subset B$ y composición de funciones. Si $B \subset \mathbb{R}$, se suele decir que f es una función ***de valores reales***.

8. Axioma de sustitución *Si $S(a, b)$ es un enunciado tal que para cada $a \in A$ se puede formar el conjunto $\{b \mid S(a, b)\}$, entonces existe una función F con dominio A tal que $F(a) = \{b \mid S(a, b)\}$ para cada $a \in A$.*

Por definición, una función F tiene un destino; por tanto el axioma 8 requiere la existencia de un conjunto B tal que $F \subset A \times B$.

Podremos recordar estos axiomas si los resumimos de una forma sugerente como sigue. El axioma de extensión nos da un criterio para la igualdad de dos conjuntos. Los axiomas de especificación, de apareamiento, de la unión y de la potencia nos permiten especificar subconjuntos y formar pares (y conjuntos finitos en general), intersecciones y uniones, y la colección de todos los subconjuntos de un conjunto dado (llamada ***conjunto potencia*** del conjunto dado). Postulamos la existencia de conjuntos infinitos. El axioma de elección asegura que podemos elegir un único elemento de cada conjunto de una colección de conjuntos no vacíos disjuntos dos a dos y formar un conjunto con dichos elementos. El axioma de sustitución indica que podemos sustituir cada elemento de un conjunto dado por un conjunto que dependa de ese elemento.

Para dar una demostración completamente detallada en matemáticas debemos recurrir a estos axiomas y a los principios elementales de la lógica. En la práctica usamos principalmente las operaciones de la teoría de conjuntos: unión, intersección, complemento, diferencia, potencia y elección (esta última casi siempre de manera implícita). ♦

Ejemplos resueltos de la introducción

Ejemplo I.1 *Demuéstrese que dados los conjuntos A, B, $C \subset S$, se cumple la **ley distributiva:***

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

[4] En el original no se incluye esta aclaración. (*N. del R.T.*)

Solución Un método es demostrar que cada miembro de la igualdad es un subconjunto del otro. Primero tomemos $x \in A \cap (B \cup C)$. Esto significa que x es un miembro de A y de $B \cup C$. Luego x está en A y además está en $A \cap B$ o en $A \cap C$; es decir, $x \in (A \cap B) \cup (A \cap C)$, así que $A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C)$. Por otro lado, si $x \in (A \cap B) \cup (A \cap C)$, estará en $A \cap B$ o en $A \cap C$. Si $x \in A \cap B$, entonces x está en A y en B; en particular, x está en A y en $B \cup C$, así que $x \in A \cap (B \cup C)$. De manera similar, si $x \in A \cap C$ concluimos que $x \in A \cap (B \cup C)$. Por lo tanto, $(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C)$, y tenemos ya la igualdad. La ley distributiva se puede verificar con un diagrama como el de la figura I-8. ♦

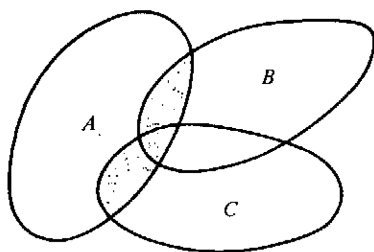


FIGURA I-8 Ley distributiva

Ejemplo I.2 *Demuéstrese que dados A, $B \subset S$,*

$$A \subset B \Leftrightarrow S\backslash A \supset S\backslash B.$$

Solución Primero probamos que $A \subset B$ implica $S\backslash B \subset S\backslash A$. Supongamos que $A \subset B$ y que $x \in S\backslash B$. Entonces, $x \notin B$ y, por lo tanto, $x \notin A$ (porque $x \in A \Rightarrow x \in B$), de donde $x \in S\backslash A$, lo que demuestra que $S\backslash B \subset S\backslash A$. Para demostrar el inverso, supongamos que $S\backslash B \subset S\backslash A$ y que $x \in A$. Entonces, $x \notin S\backslash A$ y así $x \notin S\backslash B$ (porque $S\backslash B \subset S\backslash A$), de donde $x \in B$. Por lo tanto, $A \subset B$. ♦

Ejemplo I.3 *Sea $f(x) = x^2$ (definida sobre el conjunto de todos los números reales) y $B = \{y \mid y \geq 1\}$. Calcúlese el conjunto $f^{-1}(B)$.*

Solución Por definición, $f^{-1}(B)$ consta de todos los x tales que $f(x) \in B$, es decir, todos los x tales que $x^2 \geq 1$. Esto ocurre sii $x \geq 1$ o $x \leq -1$; por lo tanto, $f^{-1}(B) = \{x \mid x \geq 1\} \cup \{x \mid x \leq -1\}$. ♦

Ejemplo I.4 *Demuéstrese que $f(A \cap B) \subset f(A) \cap f(B)$ y dése un ejemplo de A, B y f con la propiedad de que $f(A \cap B) \neq f(A) \cap f(B)$.*

Solución Si $y \in f(A \cap B)$, entonces existe algún $x \in A \cap B$ tal que $y = f(x)$. Sabemos que $x \in A$ y $x \in B$, de donde $y \in f(A)$ e $y \in f(B)$. Esto demuestra que $y \in f(A) \cap f(B)$; por lo tanto, $f(A \cap B) \subset f(A) \cap f(B)$. Para el ejemplo, sean $A = \{x \in \mathbb{Z} \mid x \geq 0\}$, $B = \{x \in \mathbb{Z} \mid x \leq 0\}$, con $f : \mathbb{Z} \to \mathbb{Z}$ definida por $f(x) = x^2$. Entonces, $f(A) = A$, $f(B) = A$ y $f(A) \cap f(B) = A$. Sin embargo, $A \cap B = \{0\}$, y, por lo tanto, $f(A \cap B) = f(\{0\}) = \{0\} \neq A$. ♦

Ejemplo I.5 *Pruébese, utilizando la inducción, que $1 + 2 + \cdots + n = n(n+1)/2$ para todo $n \in \mathbb{N}$.*

Solución Sea $P(i)$ el enunciado $1 + \cdots + i = i(i+1)/2$ cuya veracidad o falsedad estamos tratando de determinar. Verificamos las condiciones 1 y 2 del método de inducción dadas en la primera parte de esta introducción. Es obvio que $P(1)$ es cierto, porque se reduce a $1 = 1 \cdot 2/2$. Ahora *supondremos* que el enunciado $P(n)$ es cierto y *probaremos* que $P(n+1)$ es cierto:

$$1 + 2 + \cdots + n + (n+1) = (1 + 2 + \cdots + n) + (n+1)$$

$$= \frac{n \cdot (n+1)}{2} + (n+1) = \frac{n \cdot (n+1) + 2 \cdot (n+1)}{2}$$

$$= \frac{(n+1)(n+2)}{2} = \frac{(n+1)[(n+1)+1]}{2}.$$

Entonces, el enunciado es cierto para $n + 1$. ♦

Ejemplo I.6 *Sea A un conjunto y sea $\mathcal{P}(A)$ el conjunto de todos los subconjuntos de A. Pruébese que A y $\mathcal{P}(A)$ no tienen la misma cardinalidad.*

Solución Supongamos que existe una biyección $f : A \to \mathcal{P}(A)$; deduciremos entonces una contradicción. Sea $B = \{x \in A \mid x \notin f(x)\}$; existe $y \in A$ tal que $f(y) = B$, porque f es suprayectiva. Si $y \in B$, entonces por la definición de B concluimos que $y \notin B$. De manera similar, si $y \notin B$, entonces concluimos que $y \in B$; en cualquiera de los dos casos obtenemos una contradicción. En realidad, el argumento prueba que ni siquiera existe una función $f : A \to \mathcal{P}(A)$ que sea suprayectiva. ♦

Ejercicios de la introducción

1. Las funciones siguientes se definen proporcionando $f(x)$, el dominio S y el destino T. Dados $A \subset S$ y $B \subset T$, calcúlense $f(A)$ y $f^{-1}(B)$.

a. $f(x) = x^2$, $S = \{-1, 0, 1\}$, T = todos los números reales,

$A = \{-1, 1\}$, $B = \{0, 1\}$.

b. $f(x) = \begin{cases} x^2, & \text{si } x \geq 0 \\ -x^2, & \text{si } x < 0 \end{cases}$

$S = T$ = todos los números reales,

$A = \{x \in \text{ números reales} \mid x > 0\}$, $B = \{0\}$.

c. $f(x) = \begin{cases} 1, & \text{si } x > 0 \\ 0, & \text{si } x = 0 \\ -1, & \text{si } x < 0 \end{cases}$

$S = T$ = todos los números reales,

$A = B = \{x \in \text{ números reales} \mid -2 < x < 1\}$.

2. Determínese si las funciones que aparecen en el ejercicio 1 son inyectivas, suprayectivas o biyectivas para los dominios y destinos dados.

3. Sean $f: S \to T$ una función, $C_1, C_2 \subset T$ y $D_1, D_2 \subset S$. Pruébese:

a. $f^{-1}(C_1 \cup C_2) = f^{-1}(C_1) \cup f^{-1}(C_2)$.

b. $f(D_1 \cup D_2) = f(D_1) \cup f(D_2)$.

c. $f^{-1}(C_1 \cap C_2) = f^{-1}(C_1) \cap f^{-1}(C_2)$.

d. $f(D_1 \cap D_2) \subset f(D_1) \cap f(D_2)$.

4. Verifíquense las relaciones de la (a) a la (d) en el ejercicio 3 para cada una de las funciones de la (a) a la (c) del ejercicio 1, con los siguientes conjuntos: utilícense los conjuntos de la parte **a** que sigue a continuación para el ejercicio 1(a), los conjuntos de la parte **b** para el ejercicio 1(b) y los conjuntos de la parte **c** para el ejercicio 1(c):

a. C_1 = todo $x > 0$, $D_1 = \{-1, 1\}$, C_2 = todo $x \leq 0$, $D_2 = \{0, 1\}$;

b. C_1 = todo $x \geq 0$, D_1 = todo $x > 0$, C_2 = todo $x \leq 2$, D_2 = todo $x \geq -1$;

c. C_1 = todo $x \geq 0$, D_1 = todo x, C_2 = todo $x > -1$, D_2 = todo $x > 0$.

5. Si $f: S \rightarrow T$ es una función de S en T, demuéstrese que las siguientes proposiciones son equivalentes (cada una implica las otras dos):

a. f es inyectiva.

b. Para todo y en T, el conjunto $f^{-1}(\{y\})$ contiene como mucho un punto.

c. $f(D_1 \cap D_2) = f(D_1) \cap f(D_2)$ para todos los subconjuntos D_1 y D_2 de S.

Desarróllese un criterio similiar para la suprayectividad.

6. Pruébese que el conjunto de los enteros positivos $\mathbf{N} = \{1, 2, 3, \ldots\}$ tiene tantos elementos como el conjunto de los enteros, estableciendo una correspondencia biunívoca entre $\mathbf{N}$ y $\mathbf{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$. Conclúyase que $\mathbf{Z}$ es numerable.

7. Sea A un conjunto finito con N elementos y sea $\mathcal{P}(A)$ la colección de todos los subconjuntos de A, incluyendo el conjunto vacío. Pruébese que $\mathcal{P}(A)$ tiene 2^N elementos.

8. **a.** Sea $N = \{0, 1, 2, 3, \ldots\}$. Definamos $\varphi: N \times N \rightarrow N$ como $\varphi(i, j) = j + \frac{1}{2}k(k+1)$, donde $k = i + j$. Demuéstrese que φ es una biyección y que tiene algo que ver con el diagrama siguiente:

5	20					
4	14	19				
3	9	13	18			
2	5	8	12	17		
1	2	4	7	11	16	
0	0	1	3	6	10	15
		1	2	3	4	5

b. Demuéstrese que si $A_1, A_2, \ldots$ son conjuntos numerables, entonces $A_1 \cup A_2 \cup \cdots$ es numerable.

9. Sea $\mathcal{A}$ una familia de subconjuntos de un conjunto S. Escribimos $\cup\mathcal{A}$ para la unión de todos los elementos de $\mathcal{A}$ y, de manera similar, definimos $\cap\mathcal{A}$. Supongamos que $\mathcal{B} \supset \mathcal{A}$. Pruébese que $\cup\mathcal{A} \subset \cup\mathcal{B}$ y que $\cap\mathcal{B} \subset \cap\mathcal{A}$.

10. Sean $f : S \to T$, $g : T \to U$ y $h : U \to V$ funciones. Pruébese que $h \circ (f \circ g) = (h \circ f) \circ g$ (es decir, que *la composición es asociativa*).

11. Sean $f : S \to T$, $g : T \to U$ funciones dadas. Demuéstrese que para todo $C \subset U$, $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$.

12. Sea $\mathcal{A}$ una colección de subconjuntos de un conjunto S y $\mathcal{B}$ la colección de los complementarios; es decir, $B \in \mathcal{B}$ sii $S \backslash B \in \mathcal{A}$. Pruébense las ***leyes de Morgan***:

a. $S \backslash \cup \mathcal{A} = \cap \mathcal{B}$.

b. $S \backslash \cap \mathcal{A} = \cup \mathcal{B}$.

Aquí $\cup \mathcal{A}$ denota la unión de todos los conjuntos de $\mathcal{A}$.

Por ejemplo, si $\mathcal{A} = \{A_1, A_2\}$ entonces

a. es $S \backslash (A_1 \cup A_2) = (S \backslash A_1) \cap (S \backslash A_2)$ y

b. es $S \backslash (A_1 \cap A_2) = (S \backslash A_1) \cup (S \backslash A_2)$.

13. Sean $A, B \subset S$. Demuéstrese que

$$A \times B = \varnothing \Leftrightarrow A = \varnothing \text{ o } B = \varnothing.$$

14. Demuéstrese que

a. $(A \times B) \cup (A' \times B) = (A \cup A') \times B$.

b. $(A \times B) \cap (A' \times B') = (A \cap A') \times (B \cap B')$.

15. Demuéstrese que

a. $f : S \to T$ es inyectiva sii existe una función $g : T \to S$ tal que $g \circ f = I_S$; llamamos a g una ***inversa por la izquierda*** de f.

b. $f : S \to T$ es suprayectiva sii existe una función $h : T \to S$ tal que $f \circ h = I_T$; llamamos a h una ***inversa por la derecha*** de f.

c. Una función $f : S \to T$ es una biyección sii existe una función $g : T \to S$ tal que $f \circ g = I_T$ y $g \circ f = I_S$. Demuéstrese también que $g = f^{-1}$ y que está determinada de manera única.

16. Sean $f : S \to T$ y $g : T \to U$ biyecciones. Demuéstrese que $g \circ f$ es una biyección y que $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. [Sugerencia: Úsese el ejercicio 15(c).]

17. (Para este problema quizá tenga que repasar el álgebra lineal.) Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

una matriz $m \times n$ donde a_{ij} son números reales. Utilícese el ejercicio 15 para demostrar que A tiene rango m si y sólo si existe una matriz B tal que AB es la matriz identidad $m \times m$.

Capítulo 1
La recta real y el espacio euclídeo

Los objetos básicos del análisis matemático son los números y las funciones. En este libro estudiaremos la teoría de la convergencia, continuidad, diferenciación e integración de funciones de una o varias variables, tal como se ha desarrollado en los últimos 250 años. A lo largo del proceso ubicaremos esta teoría en el contexto de algunas de las técnicas del análisis moderno. Este capítulo comienza con la recta real y el espacio n-dimensional, y es indispensable para estudiar de manera precisa y entender de forma clara el cálculo de varias variables. Algunas partes de este capítulo pueden parecer un repaso de materias que tal vez ya se hayan estudiado en cursos anteriores de matemáticas. Sin embargo, intentamos cubrir los temas de una manera más rigurosa, dando algunas propiedades más de la recta real como preparación para posteriores estudios.

§1.1 Cuerpos ordenados y los sistemas numéricos

Cuerpos ordenados

Comencemos con las propiedades más importantes de los números reales. Los argumentos heurísticos (es decir, intuitivos) que comúnmente se utilizan para hablar de los números reales deben ser ya familiares para el lector. Comenzamos con los enteros no negativos 0, 1, 2, 3, . . . , agregamos los enteros negativos y los racionales no enteros. El sistema de los números reales se obtiene añadiendo a los racionales todos los límites no racionales de números racionales. Por ejemplo, el número irracional $\sqrt{2}$ se obtiene como límite de una sucesión creciente (o monótona) x_n con $x_n^2 < 2$ y x_n racional. Podemos emplear sucesiones decimales del tipo 1, 1.4, 1.41, 1.414, . . . como x_n. Es un hecho bien conocido, probado por primera vez por Euclides, que $\sqrt{2}$ no es racional (véase el ejercicio 2 al final de este capítulo).

¿Cómo se construyen los números reales de una manera formal? Este proceso es un poco largo, tedioso en algunos momentos y difícil en otros. En lo que resta de la sección daremos un esbozo del mismo. Primero resaltaremos algunas propiedades importantes que, pensamos, deben tener los números reales. La construcción de sistemas de objetos que obedecen estas propiedades a partir de la lógica y la teoría de conjuntos es un paso importante en la fundamentación del análisis, que fue dado por Karl Weierstrass, Richard Dedekind, Georg Cantor y otros, entre el final del siglo pasado y principios de éste, y, de hecho, algunos refinamientos de la idea continúan produciéndose hoy en día.

Empecemos con un conjunto dado a cuyos elementos se les llamará ***números*** y consideremos los siguientes axiomas para dicho conjunto:

Axiomas de cuerpo

I. ***Axiomas de la suma*** Existe una operación de suma " + " tal que para cualesquiera números x, y, z se cumple lo siguiente:

1. $x + y = y + x$. ***conmutatividad***

2. $x + (y + z) = (x + y) + z$. ***asociatividad***

3. Existe un número, 0, tal que $x + 0 = x$. ***elemento neutro (cero)***

4. Para cada x existe otro número, que denotaremos $-x$, tal que $x + (-x) = 0$; escribimos $y - x = y + (-x)$. ***inverso respecto de la suma***

II. ***Axiomas de la multiplicación*** Existe una operación de multiplicación "·" tal que

5. $x \cdot y = y \cdot x$. ***conmutatividad***

6. $x \cdot (y \cdot z) = (x \cdot y) \cdot z$. ***asociatividad***

7. Existe un número 1 tal que $1 \cdot x = x$. ***elemento neutro (unidad)***

8. Para cada $x \neq 0$ existe otro número, que denotaremos x^{-1}, tal que $x \cdot x^{-1} = 1$; escribimos $y \cdot x^{-1} = y/x$. ***recíproco***

9. $x \cdot (y + z) = x \cdot y + x \cdot z$. ***ley distributiva***

10. $1 \neq 0$. ***no trivialidad***

Cualquier conjunto o "sistema numérico" con operaciones + y · que obedezcan estas leyes se denomina ***cuerpo***. Por ejemplo, los números racionales son un cuerpo, pero los naturales no (porque el recíproco de un entero diferente de 1 no es un entero). A partir de este momento, escribiremos sólo xy en lugar de $x \cdot y$. La regla 10 excluye el caso trivial del cuerpo cuyo único elemento es 0.

Si comenzamos con los números naturales N, entonces se cumplen los axiomas 1, 2, 5, 6, 7, 9 y 10. La inclusión de los enteros negativos y el cero para formar Z nos da un sistema que obedece a los axiomas del 1 al 7, el 9 y el 10. Por último, el hecho de agregar los recíprocos para formar los números racionales Q conduce a un sistema que obedece todos estos axiomas. Trabajaremos sobre esta idea más adelante en esta sección.

De los axiomas de cuerpo se pueden deducir las propiedades para la manipulación de igualdades algebraicas, tal como la deducción de la identidad $(a - b)^2 = a^2 - 2ab + b^2$ y las leyes de los exponentes. Daremos unos ejemplos en la proposición **1.1.2**, así como en los ejemplos resueltos y ejercicios, y supondremos esas propiedades válidas a partir de ese momento.

Los números reales poseen también un orden natural. Habitualmente los imaginamos formando una recta.

Axiomas de orden

III. Existe una relación " $\leq$ " tal que

11.	Para cada x tenemos $x \leq x$.	***reflexividad***
12.	Si $x \leq y$ e $y \leq x$, entonces $x = y$.	***antisimetría***
13.	Si $x \leq y$ e $y \leq z$, entonces $x \leq z$.	***transitividad***
14.	Para cada par de números (x,y), $x \leq y$ o bien $y \leq x$.	***orden lineal***
15.	Si $x \leq y$, entonces $x + z \leq y + z$ para todo z.	***compatibilidad de*** $\leq$ ***y*** $+$
16.	Si $0 \leq x$ y $0 \leq y$, entonces $0 \leq xy$.	***compatibilidad de*** $\leq$ ***y*** $\cdot$

Las propiedades 11 a 13 establecen que $\leq$ es un ***orden parcial***. La propiedad 14 dice que todo par de números son comparables, lo que se describe diciendo que $\leq$ es un ***orden lineal*** u ***orden total***. Las propiedades 15 y 16 dicen que el orden y las operaciones son compatibles. Un sistema que satisface las 16 propiedades anteriores se denomina ***cuerpo ordenado***. Los números reales $\mathbb{R}$ y los racionales Q son ejemplos concretos. Por definición, $x < y$ significa que $x \leq y$ y $x \neq y$. También, $x \geq y$ significa $y \leq x$, y $x > y$ significa $y < x$. Las propiedades 11, 12 y 14 se combinan para dar lugar a la siguiente observación.

1.1.1 Ley de tricotomía *Si x e y son elementos de un cuerpo ordenado, entonces una y sólo una de las siguientes relaciones es cierta: $x < y$, $x = y$ o $x > y$.*

Existen otros sistemas, además de los números reales, en los que algunos de estos axiomas cumplen un cierto papel. Por ejemplo, los axiomas del 1 al 4 definen un ***grupo conmutativo***. Los axiomas del 1 al 9, excluyendo los axiomas 5 y 8, definen un ***anillo***; este concepto es útil, por ejemplo, para estudiar operaciones algebraicas sobre el conjunto de las matrices $n \times n$. Obsérvese que el conjunto de todos los subconjuntos de un conjunto dado S, donde por $A \leq B$ entendemos $A \subset B$, forma un conjunto parcialmente ordenado, y que este orden no satisface la propiedad 14.

1.1.2 Proposición

En un cuerpo ordenado son válidas las siguientes propiedades:

i. ***Unicidad de los neutros*** *Si* $a + x = a$ *para todo* a, *entonces* $x = 0$. *Si* $a \cdot x = a$ *para todo* a, *entonces* $x = 1$.

ii. ***Unicidad de los inversos*** *Si* $a + x = 0$, entonces $x = -a$. *Si* $ax = 1$, *entonces* $x = a^{-1}$.

iii. ***No hay divisores de cero*** *Si* $xy = 0$, *entonces* $x = 0$ o $y = 0$.

iv. ***Reglas de cancelación para la suma*** *Si* $a + x = b + x$, *entonces* $a = b$. *Si* $a + x \leq b + x$, *entonces* $a \leq b$.

v. ***Regla de cancelación para la multiplicación*** *Si* $ax = bx$ *y* $x \neq 0$, *entonces* $a = b$. *Si* $ax \geq bx$ *y* $x > 0$, *entonces* $a \geq b$.

vi. $0 \cdot x = 0$ *para todo* x.

vii. $-(-x) = x$ *para todo* x.

viii. $-x = (-1)x$ *para todo* x.

ix. *Si* $x \neq 0$, *entonces* $x^{-1} \neq 0$ y $(x^{-1})^{-1} = x$.

x. *Si* $x \neq 0$ *e* $y \neq 0$, *entonces* $xy \neq 0$ *y* $(xy)^{-1} = x^{-1}y^{-1}$.

xi. *Si* $x \leq y$ y $0 \leq z$, *entonces* $xz \leq yz$. *Si* $x \leq y$ y $z \leq 0$, *entonces* $yz \leq xz$.

xii. *Si* $x \leq 0$ e $y \leq 0$, *entonces* $xy \geq 0$. *Si* $x \leq 0$ e $y \geq 0$, *entonces* $xy \leq 0$.

xiii. $0 < 1$.

xiv. $x^2 \geq 0$ *para todo* x.

Con el propósito de ilustrar las ideas de las demostraciones de **ii, vi, viii, xiii** y **xiv**, éstas se dan al final del capítulo.

El propósito de los axiomas de un cuerpo ordenado es aislar las propiedades básicas que necesitamos para manejar las igualdades y desigualdades algebraicas. Por ejem-

plo, he aquí un caso típico: por **1.1.2xiv**, para cualesquiera números a y b, $(a-b)^2 \geq 0$; entonces, $a^2 - 2ab + b^2 \geq 0$, y, por lo tanto,

$$ab \leq \frac{1}{2}(a^2 + b^2)$$

para cualquier par de números a, b. Este tipo de manipulaciones es importante en análisis y volverán a aparecer más adelante.

1.1.3 Ejemplo *Usando los axiomas y propiedades de cuerpo ordenado dados en esta sección, pruébese que*

$$a^2 - b^2 = (a-b)(a+b).$$

Solución Por la ley distributiva,

$$(a-b)(a+b) = (a-b)\cdot a + (a-b)\cdot b.$$

Ahora usamos las leyes conmutativa y distributiva otra vez, junto con $a - b = a + (-b)$:

$$(a-b)\cdot a + (a-b)\cdot b = a\cdot(a-b) + b\cdot(a-b) = a^2 + a\cdot(-b) + b\cdot a + b\cdot(-b).$$

Ahora, $a\cdot(-b) = a\cdot(-1)\cdot b = (-1)ab = -(ab)$ por **1.1.2viii**, la asociatividad y la conmutatividad. De manera similar, $b\cdot(-b) = -b^2$. Entonces, $(a-b)(a+b)$ es igual a

$$\begin{aligned} a^2 + a\cdot(-b) + b\cdot a + b\cdot(-b) &= a^2 - (ab) + ba - b^2 \\ &= a^2 - ab + ab + b^2 \text{ (por el axioma 5)} \\ &= a^2 - b^2 \text{ (por los axiomas 3 y 4).} \end{aligned}$$ ♦

1.1.4 Ejemplo *En un cuerpo ordenado, pruébese que si $0 \leq x < y$, entonces $x^2 < y^2$.*

Solución Si $0 \leq x < y$, entonces $0 \leq x \leq y$, y por **1.1.2xi**, $x^2 \leq yx$. Por el mismo razonamiento, $x \leq y$ implica $xy \leq y^2$. Por lo tanto,

$$\begin{aligned} x^2 &\leq yx = xy \text{ (por la conmutatividad)} \\ &\leq y^2, \end{aligned}$$

de donde $x^2 \leq y^2$. Sólo queda excluir la posibilidad de que x^2 sea igual a y^2. Pero si $x^2 = y^2$ entonces

$$\begin{aligned} x^2 - y^2 &= 0 \qquad \text{(sume } -y^2 \text{ a cada miembro)} \\ (x-y)(x+y) &= 0 \qquad \text{(por el ejemplo } \mathbf{1.1.3}\text{)} \end{aligned}$$

Por **1.1.2xii** tenemos $0 \leq x$ e $y > 0$. Ahora bien, $x + y \neq 0$, ya que $x + y = 0$ implicaría que $y = (-x) \leq 0$, de modo que $y \leq 0$, lo que es imposible por la ley de tricotomía. Por la cancelación para la multiplicación,

$$(x - y)(x + y) = 0$$

implica que $x - y = 0$, es decir, que $x = y$. Pero sabemos que $x < y$, luego este caso está excluido, como se deseaba. ♦

Podemos introducir otros símbolos familiares. El ***módulo*** o ***valor absoluto*** de un número x se denota $|x|$ y en un cuerpo ordenado se define como $|x| = x$ si $x \geq 0$ y $|x| = -x$ si $x < 0$. La ***distancia*** entre dos elementos x e y es $|x - y|$. Vale la pena resaltar ciertas propiedades del valor absoluto y de la distancia, tanto por su propia importancia como porque las generalizaremos para otras situaciones más adelante. Sus demostraciones son inmediatas, así que las omitiremos.

1.1.5 Proposición

i. $|x| \geq 0$ *para todo* x.

ii. $|x| = 0$ *si y sólo si* $x = 0$.

iii. $|xy| = |x|\,|y|$.

iv. $|x + y| \leq |x| + |y|$. ***desigualdad triangular***

v. $||x| - |y|| \leq |x - y|$. ***desigualdad triangular alternativa***

Por ejemplo, para probar **iv** se consideran tres casos; así, si $x \geq 0$, $y \leq 0$ y $x + y \geq 0$, entonces $|x + y| = x + y \leq x + y - y = x = |x| \leq |x| + |y|$; los demás casos son análogos.

1.1.6 Proposición *Si $d(x,y) = x - y$, entonces*

i. $d(x, y) \geq 0$.

ii. $d(x, y) = 0$ *si y sólo si* $x = y$.

iii. $d(x, y) = d(y, x)$.

iv. $d(x, y) \leq d(x, z) + d(z, y)$. ***desigualdad triangular***

El término *desigualdad triangular* proviene de la desigualdad correspondiente cuando x e y no son números, sino puntos en el plano y $d(x, y)$ es la distancia euclídea usual

entre ellos. En este caso, la desigualdad expresa el hecho conocido de que la suma de las longitudes de dos lados de un triángulo siempre es igual al menos a la longitud del tercer lado. Estudiaremos esta situación más general en §**1.6**.

Los sistemas numéricos

Repasemos brevemente los diferentes sistemas numéricos para recordarlos y establecer la notación.

Los números naturales

Los ***números naturales*** son los que usamos para contar N = {1, 2, 3, 4, . . .}. Si agregamos el cero obtenemos el conjunto **N** = {0, 1, 2, 3, . . .} de los enteros no negativos. Los elementos neutros 0 y 1 están en **N**, y la suma y la multiplicación están definidas; sin embargo, no tenemos negativos ni recíprocos. Con el orden usual, **N** satisface todos los axiomas del 1 al 16 excepto el 4 y el 8. La propiedad más notable de **N** es que sus elementos son los números que usamos para contar. Hay un primer elemento, 0, y para cada uno de sus elementos existe un "siguiente" bien definido, que se obtiene sumándole 1. Repitiendo el proceso de sumar 1 se genera todo **N**. Esto se formaliza mediante el

Principio de inducción matemática *Si S es un subconjunto de* **N** *tal que* $0 \in S$ *y* $k + 1$ *está en S siempre que k está en S, entonces* $S =$ **N**.

Como ya vimos en el capítulo de introducción, ésta es una herramienta muy poderosa, no sólo en la teoría de números sino en toda la matemática. Vale la pena puntualizar algunas consecuencias de este principio. En efecto, el conjunto **N** tiene un elemento mínimo: el 0. También tiene un elemento mínimo cualquier subconjunto no vacío de **N**. Decimos que N está ***bien ordenado*** por $\leq$ o que satisface la propiedad de buena ordenación. También es común referirse a $\leq$ como una buena ordenación de **N**.

1.1.7 Proposición $\leq$ *es una buena ordenación de* **N**; *es decir,* **N** *tiene la propiedad de buena ordenación*:

Propiedad de buena ordenación *Si S es un subconjunto no vacío de* **N**, *entonces S tiene un elemento mínimo; es decir, existe $s_0 \in S$ tal que $s_0 \leq x$ para todo x en S.*

La demostración, dada al final del capítulo, ilustra una manera de usar la inducción que puede ser nueva para el lector. Aunque es necesaria una demostración formal, la validez de la propiedad de buena ordenación es intuitivamente clara.

Como consecuencia de la propiedad de buena ordenación, se puede demostrar que los números naturales también satisfacen el

Principio de inducción completa *Si $S \subset \mathbf{N}$ es un subconjunto tal que $\{x \in \mathbf{N} \mid x < n\} \subset S$ implica $n \in S$ y, por lo tanto, $S = \mathbf{N}$.*

Esta forma del principio de inducción se puede aplicar a otros conjuntos bien ordenados además de **N**. No equivale al principio de inducción finita; es decir, existen conjuntos bien ordenados que satisfacen el principio de inducción completa pero no el principio de inducción matemática.

Los enteros

Para pasar de **N** a los ***enteros***, $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$, agregamos los enteros negativos, con lo que aseguramos la existencia de soluciones a toda ecuación de la forma $x + n = m$ con n y m en **N**. Al hacer esto hemos producido un sistema en el cual existe una solución única $x = m - n$ siempre que n y m estén en $\mathbb{Z}$. Este sistema satisface los axiomas del 1 al 16 excepto el 8. Como se mencionó antes, una estructura que satisface del 1 al 4, el 6 y el 9 se llama ***anillo***. Con el 5 es un ***anillo conmutativo***, y con el 7 es un ***anillo con unidad*** o ***anillo con identidad***. Hay ejemplos de anillos conmutativos además de $\mathbb{Z}$ (piénsese en la "aritmética del reloj" de la escuela elemental). Cuando exigimos los axiomas del 10 al 16, las posibilidades se reducen. En cierto sentido, el sistema más sencillo posible es $\mathbb{Z}$. Al extender **N** en ambas direcciones, hemos perdido el principio de inducción matemática, pero de alguna manera lo tenemos en dos direcciones: si S es un subconjunto de $\mathbb{Z}$, tal que $0 \in S$ y $k + 1$ y $k - 1$ están en S siempre que k esté en S, entonces $S = \mathbb{Z}$.

Los números racionales

El conjunto $\mathbb{Q}$ de los ***números racionales*** consta de todos aquellos números que se pueden expresar como fracciones o cocientes de enteros: m/n con $n \neq 0$. Podemos verlo como una extensión de $\mathbb{Z}$ para que incluya las soluciones de todas las ecuaciones de la forma $nx + m = 0$ con $n, m \in \mathbb{Z}$ y $n \neq 0$. En este proceso obtenemos un sistema, $\mathbb{Q}$, en el cual toda ecuación $ax + b = 0$ con a y b en $\mathbb{Q}$ y $a \neq 0$ tiene la solución única $x = -a^{-1} \cdot b$. Un sistema como éste, que satisface 1–10, se denomina ***cuerpo***. Ejemplos de cuerpos son $\mathbb{Q}$ y los números reales $\mathbb{R}$. En §**1.8** estudiaremos el cuerpo de los números

complejos $\mathbb{C}$, pero hay otros más, incluyendo algunos con un número finito de elementos. Debemos excluir estos últimos y $\mathbb{C}$ si queremos un cuerpo ordenado que satisfaga también 11–16. Pero aún quedan $\mathbb{R}$, $\mathbb{Q}$ y otros. Si se impone algún tipo de condiciones de minimalidad, como pedir que ningún subconjunto propio sea también un cuerpo, entonces se puede demostrar que $\mathbb{Q}$ es esencialmente la única posibilidad.

Continuamos con dos proposiciones: la primera dice que $\mathbb{Q}$ tiene muchos elementos; la segunda, que no demasiados.

1.1.8 Proposición *$\mathbb{Q}$ es "denso en sí mismo": si x e y están en $\mathbb{Q}$, entonces existe un elemento z en $\mathbb{Q}$ tal que $x < z < y$.*

Esta proposición se puede verificar eligiendo $z = (x + y)/2$.

Si comenzamos a marcar los números racionales de la recta numérica, encontraremos que se distribuyen de una manera densa en la misma hasta un punto en que parece que la llenan (sabemos que no; por ejemplo, falta $\sqrt{2}$).

1.1.9 Proposición *$\mathbb{Q}$ es numerable.*

Esto significa que $\mathbb{Q}$ se puede poner en correspondencia biunívoca con $\mathbb{N}$ (o un subconjunto de éste). El conjunto $\mathbb{Q}$ en su totalidad se puede presentar en una lista o sucesión. Una manera de hacer esto es:

$$0, 1, -1, 2, -2, \frac{1}{2}, -\frac{1}{2}, 3, -3, \frac{1}{3}, \frac{2}{3}, -\frac{1}{3}, -\frac{2}{3}, 4, -4, \frac{1}{4}, \frac{3}{4}, -\frac{1}{4}, -\frac{3}{4}, 5, \ldots$$

Es decir, empezamos con 0, 1, –1, 2, –2, luego siguen aquellas fracciones que tienen denominador 2, expresadas en su mínima expresión, luego 3 y –3 y aquellas fracciones con denominador 3, etcétera, evitando repeticiones.

La última propiedad de $\mathbb{Q}$ es bastante trivial, pero es muy importante en $\mathbb{R}$. Es la

Propiedad arquimediana *Si $x \in \mathbb{Q}$ entonces existe un entero n tal que $x < n$.*

Si $x \leq 0$, tómese $n = 1$. Si $x = p/q$ con $p > 0$ y $q > 0$, tómese $n = p + 1$.

En cualquier cuerpo ordenado hay una copia de $\mathbb{N}$ que se obtiene identificando n con $1 + 1 + \cdots + 1$ (n términos). Esto se extiende a una copia de $\mathbb{Z}$ incluyendo $-n$ y a una copia de $\mathbb{Q}$, identificando k/n con $n^{-1} \cdot k$ si $n \neq 0$. Un cuerpo $\mathbb{F}$ se denomina ***cuerpo***

ordenado arquimediano si encierra la propiedad arquimediana. He aquí varias formulaciones equivalentes:

1. Si $x \in$ F, entonces existe un entero n tal que $x < n$.
2. Si x e y están en F, con $0 < x < y$, entonces existe un entero k tal que $kx > y$.
3. Si $x \in$ F y $x > 0$, entonces existe un entero $n > 0$ tal que $0 < 1/n < x$.

(Para obtener la tercera caracterización, se define $y = 1/x$ y se elige $n > y$.) Así pues, Q es un ejemplo de cuerpo arquimediano ordenado. Hay otros, y hay cuerpos ordenados que no son arquimedianos.

Hay otras tareas, que, en principio, se deben realizar antes de continuar; por ejemplo, es necesario formular las leyes de los exponentes enteros. Esperamos, sin embargo, haber dado ya una idea suficientemente clara del tipo de técnicas que se pueden emplear, de tal manera que lo que queda sea creíble y podamos continuar. (El trabajo de desarrollar completamente estos temas pertenece a los cursos de álgebra.)

1.1.10 Ejemplo *Encuéntrese un ejemplo de anillo conmutativo sin identidad.*

Solución Debemos satisfacer los axiomas del 1 al 6 y el 9, pero el axioma 7 debe fallar. Tal vez el sistema más sencillo con dicha propiedad es el de los enteros pares con las operaciones normales de suma y multiplicación. ♦

1.1.11 Ejemplo *Utilícese la inducción para probar que* $1/2^n < 1/n$ *para todo entero* $n > 0$.

Solución Esto es cierto para $n = 1$ porque $1/2 < 1$. Supongamos que es válido para un entero dado $n > 1$. Entonces

$$\frac{1}{2^{n+1}} = \frac{1}{2^n \cdot 2} < \frac{1}{2n} = \frac{1}{n+n} < \frac{1}{n+1}.$$

Por tanto, es verdadero para $n + 1$. ♦

Nota. Se tiene la intención de que los ejercicios de cada sección sean relativamente rutinarios, con objeto de crear confianza y procurar un poco de práctica antes de tener que ver con los problemas del final del capítulo, que representan un reto mayor. El estudiante encontrará una combinación de problemas sencillos y difíciles al final del capítulo.

Ejercicios de §1.1

1. En un cuerpo ordenado, pruébese que $(a + b)^2 = a^2 + 2ab + b^2$.

2. En un cuerpo, demuéstrese que si $b \neq 0$ y $d \neq 0$, entonces

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}.$$

3. En un cuerpo ordenado, si $a > b$, demuéstrese que $a^2b < ab^2 + (a^3 - b^3)/3$.

4. Pruébese que en un cuerpo ordenado, si $\sqrt{2}$ es un número positivo cuyo cuadrado es 2, entonces $\sqrt{2} < 3/2$ (hágase sin llevar a cabo una evaluación numérica).

5. Encuéntrese un ejemplo de cuerpo con sólo tres elementos. Pruébese que no puede ser un cuerpo ordenado.

§1.2 La completitud y el sistema de los números reales

Los axiomas de cuerpo ordenado no son suficientes para caracterizar de forma única los números reales, porque los racionales también satisfacen esos axiomas. Necesitamos, pues, otra condición que nos asegure que los límites de racionales estén incluidos en el sistema. Esta condición se conoce como ***completitud***. El sistema de los números racionales es muy rico, pero no lo suficiente para el análisis clásico. Hay números que en cierto sentido deberían estar ahí, pero no están; por ejemplo, no existe la raíz cuadrada racional de 2, o, de manera más precisa, no hay enteros p y q tales que $(p/q)^2 = 2$. Éste es un problema no sólo para el análisis, sino también para la geometría. Fue en este contexto donde esta dificultad llamó por primera vez la atención de los matemáticos y filósofos pitagóricos de la antigua Grecia. No existe un número racional que corresponda a la longitud de la diagonal de un cuadrado cuyos lados midan 1.

Hay varias formas de introducir la propiedad de completitud. Nosotros usaremos el método de las *sucesiones monótonas* porque es el más cercano a nuestra intuición de construir números reales, como $\sqrt{2} = 1.414213\ldots$, en forma de límite de una sucesión que lo aproxima. (En la sección siguiente usaremos esta propiedad para verificar la propiedad del supremo, que se puede utilizar alternativamente como axioma.) Con objeto de plantear este axioma es preciso entrar en el tema de las sucesiones.

La idea de aproximación es central en la teoría, técnicas y aplicaciones del análisis. Una versión de esta idea es que, aun si no se puede resolver un problema de manera exacta, quizá se pueda obtener un procedimiento que, aplicado reiteradamente, nos dé una aproximación cada vez mejor. En cierto sentido, esto ocurre hasta en la aritmética. Si le pedimos a un niño que divida 1 entre 3 para obtener la expansión decimal de $1/3$, el método usual de división nunca pararía, sino que produciría una sucesión de aproximaciones cada vez mejores: 0.3, 0.33, 0.333, 0.3333, . . . Comenzaremos nuestro estu-

dio con sucesiones de números reales y luego pasaremos a sucesiones de vectores. A lo largo del libro volveremos sobre este tema repetidas veces, al estudiar sucesiones de funciones y diversas ideas sobre convergencia.

Recordemos que una ***sucesión*** en un conjunto S es simplemente una función $f: \mathbb{N} \to S$. Lo esencial de la idea es que una sucesión es una lista numerable de elementos de S colocados en un orden particular. La notación con subíndices suele usarse para etiquetar los elementos $f(n) = x_n$, aunque en ocasiones se utilizan otras formas, como $x^{(n)}$ o $x(n)$. La sucesión se dispone como una lista $x_1, x_2, x_3, \ldots$. Como el etiquetado es arbitrario, podemos hacerlo de otra forma, sin pérdida de información, tomando $y_k = x_{k+1}$ para obtener $y_0, y_1, y_2, y_3, \ldots$, lo que puede simplificar la expresión de f. Por ejemplo, la sucesión

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots$$

se puede representar como $x_n = 1/2^n$ para $n = 1, 2, 3, \ldots$ o $y_k = 1/2^{k+1}$ para $k = 0, 1, 2, 3, \ldots$. En general, escribimos una sucesión en la forma x_n.

Decimos que una sucesión *converge* a un límite x o que es convergente a x si podemos garantizar que los puntos de la sucesión estén tan cerca como deseemos de x con sólo tomar el índice suficientemente grande. Esto está expresado de forma precisa en la siguiente definición.

1.2.1 Definición *Decimos que una sucesión x_n* ***converge*** *a un* ***límite*** *x si para todo $\varepsilon > 0$ existe un entero N tal que $|x_n - x| < \varepsilon$ siempre que $n \geq N$. (El número N puede depender de ε; un ε más pequeño puede requerir un N más grande.) En ese caso escribimos* $\lim_{n\to\infty} x_n = x$ o bien $x_n \to x$ *cuando $n \to \infty$.*

La definición exige que todos los términos de la sucesión con índice mayor que N disten de x menos que ε; no es suficiente con encontrar un término cerca de x. Podemos pensar que la letra griega ε (épsilon) es una cota al error máximo permitido en la aproximación, de modo que debemos poder asegurar que, más allá de alguna N específica, todos los términos de la sucesión aproximan x con este grado de precisión. Podemos imaginarnos mejor el concepto si dibujamos los puntos sobre una recta numérica, como en la figura 1.2-1.

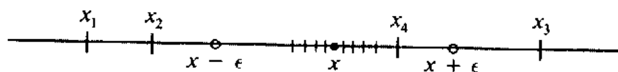


FIGURA 1.2-1 Los elementos de una sucesión x_n están a una distancia menor que ε del límite para n grande

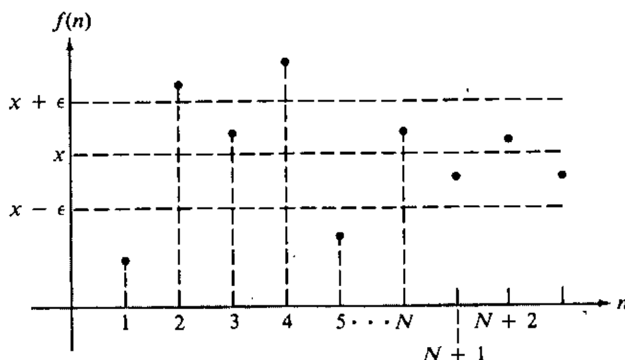


FIGURA 1.2-2 La convergencia vista en términos de una gráfica

La definición **1.2.1** exige que todos los términos, excepto quizá un número finito de ellos, estén entre $x - \varepsilon$ y $x + \varepsilon$. Podemos hacernos una mejor idea mediante la gráfica de la función que define la sucesión (figura 1.2-2). Todos los puntos que están suficientemente a la derecha deben estar en la banda de ancho 2ε con centro en la recta horizontal de altura x. *No* es suficiente que $(3, x_3)$ esté en la banda, porque $(4, x_4)$ no está.

1.2.2 Lema del sandwich *Supongamos que $x_n \to L$, $y_n \to L$ y que $x_n \leq z_n \leq y_n$ para todo n (es suficiente suponer que existe N_0 tal que $x_n \leq z_n \leq y_n$ siempre que $n > N_0$).* Entonces $z_n \to L$.

Intuitivamente ha se ser así, porque z_n está acorralado por arriba y por abajo por números que se acercan a L. La demostración completa se da al final del capítulo. Veremos más variaciones sobre este tema más adelante. Otro hecho que conviene ver es que el límite no puede exceder las cotas de la sucesión:

1.2.3 Proposición *Si $a \leq x_n \leq b$ para todo n y $x_n \to x$, entonces $a \leq x \leq b$.*

La demostración, que dejamos al lector, puede ser inmediata, utilizando un método similar al usado en el lema del sandwich.

1.2.4 Ejemplo *La sucesión $x_n = (-1)^n$, $n = 1, 2, \ldots$, ilustra varias cosas:*

1. Los puntos de una sucesión se pueden repetir; es decir, no es necesario que la función de $\mathbb{N}$ que la define sea inyectiva.

2. La sucesión no converge. Sea $\varepsilon = 1/2$. Entonces sea cual fuere el límite x que se propusiera para la sucesión, habría términos en la sucesión que distarían de x más que ε. Para verlo, supongamos primero que $x \geq 0$; entonces $|x_n - x| = |-1 - x| \geq 1 > \varepsilon$ para todo n impar. Si $x < 0$ entonces $|x_n - x| = |1 - x| \geq 1 > \varepsilon$ para todo n par.
3. La sucesión se acumula con frecuencia en el 1 y en el -1. Si sólo usamos los términos pares, convergen a 1, mientras que si usamos sólo los términos impares, obtenemos el límite -1. Esto ilustra las ideas de punto de límite y subsucesión que se estudiarán en §**1.5.** ♦

La siguiente proposición se suele resumir con la frase "El límite es único", un resultado perfectamente plausible. (Véase, de nuevo, el final del capítulo para la demostración.)

1.2.5 Proposición *Si x_n es una sucesión en un cuerpo ordenado tal que $x_n \to x$ y $x_n \to y$, entonces $x = y$.*

Hagamos notar un hecho sencillo pero importante antes de desviarnos hacia la aritmética de las sucesiones. Decimos que una sucesión está ***acotada*** si el conjunto de puntos que la forman está acotado, es decir, si existe un número M tal que $|x_n| \leq M$ para todo n. Decimos que está ***acotada superiormente*** si existe un número B tal que $x_n \leq B$ para todo n, y ***acotada inferiormente*** si existe un número A tal que $A \leq x_n$ para todo n. Claramente, una sucesión está acotada si y sólo si lo está superior e inferiormente. Llamamos a B una ***cota superior*** y a A una ***cota inferior*** de la sucesión.

1.2.6 Proposición *Una sucesión convergente está acotada.*

Esto ocurre porque todos los elementos de la sucesión se acaban amontonando cerca del límite y lo que pase con los primeros no altera el estar acotada.

Supongamos que x_n e y_n son sucesiones de un cuerpo ordenado. Las podemos sumar, restar, multiplicar y dividir término a término:

$$(x_n) + (y_n) = (x_n + y_n)$$
$$\lambda(x_n) = (\lambda x_n)$$
$$(x_n)(y_n) = (x_n y_n)$$
$$(x_n)/(y_n) = (x_n/y_n).$$

En la última operación se supone que $y_n \neq 0$. Estas operaciones se comportan más o menos como es de esperar con respecto a los límites. Si $x_n \to x$ e $y_n \to y$, entonces para

n suficientemente grande, los puntos x_n están cerca de x y los y_n están cerca de y; esperamos, pues, que $x_n + y_n$, λx_n y $x_n y_n$ estén cerca de $x + y$, λx y xy, y que si $y \neq 0$, entonces x_n/y_n esté cerca de x/y. Resumimos los resultados esperados de la siguiente manera.

1.2.7 Teorema del límite para sucesiones *Supongamos que $x_n \to x$, $y_n \to y$ y λ es constante. Entonces*

i. $x_n + y_n \to x + y$.

ii. $\lambda x_n \to \lambda x$.

iii. $x_n y_n \to xy$.

iv. *Si $y_n \neq 0$ e $y \neq 0$, entonces $x_n/y_n \to x/y$.*

Decimos que una sucesión x_n es ***creciente*** (o mejor, ***no decreciente***) si $x_n \leq x_{n+1}$ para todo n, es decir, $x_1 \leq x_2 \leq x_3 \leq \ldots$; es ***estrictamente creciente*** si $x_1 < x_2 < x_3 < \ldots$; es ***decreciente*** (o mejor, ***no creciente***) si $x_n \geq x_{n+1}$ para todo n, es decir, $x_1 \geq x_2 \geq x_3$, y es ***estrictamente decreciente*** si $x_1 > x_2 > x_3 > \ldots$. Se denomina ***monótona*** (***monótona creciente, monótona decreciente***) si está en alguno de estos casos. Una sucesión que es monótona creciente pero acotada superiormente no se puede ir a infinito y sus términos se deben amontonar cerca de algún límite, tal y como ocurre con la sucesión 1, 1.4, 1.41, 1.414, . . . , que crece convergiendo a $\sqrt{2}$.

1.2.8 Propiedad de la sucesión monótona *Sea* F *un cuerpo ordenado. Decimos que* F *tiene la* ***propiedad de la sucesión monótona*** *si toda sucesión monótona creciente acotada superiormente converge.*

1.2.9 Propiedad de completitud *Decimos que un cuerpo ordenado es* ***completo*** *si obedece la propiedad de la sucesión monótona.*

Según estas definiciones, las propiedades de la sucesión monótona y de completitud son sinónimas. Hemos usado dos nombres porque, como veremos en §**1.3** y §**1.4**, hay otras propiedades que se pueden usar para caracterizar la completitud y, en ese contexto, las distinciones en la terminología son importantes.

Como con $\sqrt{2}$, podemos usar la propiedad de la sucesión monótona para probar la existencia de raíces cuadradas de números positivos en un cuerpo ordenado completo. Dado $y > 0$, definimos una sucesión x_n tomando $x_n = N_n/2^n$, donde N_n es el entero más grande para el cual $x_n^2 \leq y$. Se puede probar que x_n es una sucesión monótona acotada superiormente y, por lo tanto, converge a un número x (véase el ejemplo **1.2.10**, a

continuación, para estudiar las técnicas empleadas). El número x satisface $x^2 = y$ y se denota $\sqrt{y}$. Aceptaremos este hecho y otros similares sobre las leyes de los exponentes (y las leyes asociadas de los logaritmos) en lo que resta del libro.

1.2.10 Ejemplo *En un cuerpo ordenado completo, definimos x_n de forma inductiva mediante $x_0 = 0, x_1 = \sqrt{2}, x_2 = \sqrt{2+x_1}, \ldots, x_n = \sqrt{2+x_{n-1}}, \ldots$ Demuéstrese que x_n converge.*

Solución Probaremos que x_n es creciente y está acotada superiormente y esto demostrará la afirmación. Primero utilicemos la inducción para probar que $x_n \geq 0$ y que $r_n = x_{n+1} - x_n \geq 0$. Claramente es cierto para $n = 0$. Supongamos que es cierto para $n - 1$; entonces $x_n = \sqrt{2+x_{n-1}} \geq \sqrt{2} > 0$ y

$$r_n = x_{n+1} - x_n = \sqrt{2+x_n} - \sqrt{2+x_{n-1}} = \frac{x_n - x_{n-1}}{\sqrt{2+x_n} + \sqrt{2+x_{n-1}}}$$
$$= \frac{r_{n-1}}{\sqrt{2+x_n} + \sqrt{2+x_{n-1}}},$$

así, $r_{n-1} \geq 0$ implica $r_n \geq 0$ y, por lo tanto, x_n es creciente. Ahora queremos probar que x_n está acotada superiormente; por ejemplo, podemos probar por inducción que $x_n \leq 2$. Está claro que $x_0, x_1 \leq 2$. Supongamos que $x_{n-1} \leq 2$; entonces

$$x_n = \sqrt{2+x_{n-1}} \leq \sqrt{2+2} \leq \sqrt{4} = 2,$$

de donde x_n es creciente y está acotada superiormente y, por lo tanto, converge. (Calcularemos el valor del límite en el ejercicio 1 de esta sección.) ♦

Como en este ejemplo, la completitud no proporciona el valor del límite; sólo nos dice que existe. Este tipo de resultados son comunes en matemáticas y se llaman ***teoremas de existencia***. Pueden parecer de poca ayuda práctica, pero en ocasiones sí son de ayuda. Una vez que sabemos que el límite existe y que tiene un valor definido podemos asignarle un símbolo, tal vez encontrar una ecuación para éste, y finalmente resolverla. Sin saber que la solución existe, dicho procedimiento nos podría dar una solución espuria.

1.2.11 Proposición *Los cuerpos ordenados completos son arquimedianos.*

¡Quizá resulte sorprendente que esto no sea obvio, pero el hecho es que existen cuerpos ordenados que *no son arquimedianos*! Un ejemplo aparece en el ejercicio 47, al final de este capítulo.

Un resultado clave que se usa en muchos otros razonamientos es el hecho "casi obvio" de que $1/n$ tiende a 0 cuando n. Esto es consecuencia de la propiedad arquimediana:

1.2.12 Ejemplo *Demuéstrese que en un cuerpo ordenado completo* $1/n \to 0$ cuando $n \to \infty$.

Solución Según la definición, dado un número $\varepsilon > 0$, debemos probar que existe un entero N con la propiedad de que si $n \geq N$, entonces $|1/n - 0| < \varepsilon$. Esto ocurre siempre que $1/N < \varepsilon$ y, por lo tanto, es suficiente elegir $N > 1/\varepsilon$, lo cual es posible por la propiedad arquimediana. ♦

El siguiente ejemplo es un poco más complicado, pero utiliza esencialmente la misma idea.

1.2.13 Ejemplo *Demuéstrese que cuando* $\sqrt{n^2+1}/n! \to 0$, $n \to \infty$.

Solución Para demostrar que $\sqrt{n^2+1}/n!$ se hace pequeño cuando n toma valores grandes, hacemos la siguiente acotación:

$$0 \leq \frac{\sqrt{n^2+1}}{n!} \leq \frac{\sqrt{2n^2}}{n!} = \frac{\sqrt{2}n}{n!} = \frac{\sqrt{2}}{(n-1!)} \leq \frac{\sqrt{2}}{n-1}.$$

Por tanto, dado que $\varepsilon > 0$ basta elegir N tal que $N > 1 + \sqrt{2}/\varepsilon$. Entonces, $n \geq N$ implica

$$0 \leq \frac{\sqrt{n^2+1}}{n!} \leq \frac{\sqrt{2}}{n-1} \leq \frac{\sqrt{2}}{N-1} < \varepsilon.$$

Esto demuestra la afirmación. ♦

Un tercer ejemplo muestra otros trucos.

1.2.14 Ejemplo *Demuéstrese que* $\lim_{n\to\infty} 1/2^n = 0$.

Método 1 Echando mano de algunos conocimientos, podemos fijar un $\varepsilon > 0$ como el error permitido y ver cuán grande debe ser N para garantizar $|1/2^n - 0| < \varepsilon$ siempre que $n \geq N$. Para hacer esto calculamos:

$$\left|\frac{1}{2^n} - 0\right| < \varepsilon \Leftrightarrow \frac{1}{2^n} < \varepsilon \Leftrightarrow \frac{1}{\varepsilon} < 2^n$$
$$\Leftrightarrow \log\frac{1}{\varepsilon} < n\log 2 \Leftrightarrow n > \frac{\log(1/\varepsilon)}{\log 2}.$$

Eligiendo cualquier $N > \log(1/\varepsilon)/\log 2$ obtenemos el resultado. Obsérvese que es importante que cada paso del cálculo sea reversible, ya que en realidad necesitamos que $n \geq N$ implique $|1/2^n - 0| < \varepsilon$. El lector exigente puede objetar, con razón, que esta demostración requiere el uso de logaritmos, los cuales no se han definido aún de manera rigurosa. He aquí otro método para esos lectores.

Método 2 Este método se basa en un resultado anterior e ilustra un truco útil: vamos a comparar nuestra sucesión con sucesiones conocidas. Por el ejemplo **1.1.11** sabemos que $0 < 1/2^n < 1/n$ para todo $n > 0$. También sabemos por **1.2.12** que $1/n \to 0$. Dado que $\varepsilon > 0$, sea N lo suficientemente grande como para que $1/n < \varepsilon$ siempre que $n \geq N$. Entonces $n \geq N$ implica $0 < 1/2^n < 1/n < \varepsilon$, o sea, $|1/2^n - 0| < \varepsilon$.

Obsérvese que el valor de N obtenido con este método es $1/\varepsilon$. Este valor podría ser más grande que el N óptimo, obtenido mediante el método 1. Para $\varepsilon = 0.001$ hemos obtenido aquí que $N = 1000$ (o mejor, 1001) funciona. El primer método muestra que cualquier N más grande que $(\log 1000)/\log 2 \approx 9.97$ funcionaría. Tenemos un pésimo valor de N a cambio de un argumento más sencillo. Desde un punto de vista teórico, a menudo es una buena idea usar estas estimaciones. Desde un punto de vista computacional, el esfuerzo de tener estimaciones óptimas puede valer la pena. El segundo aspecto que debe tenerse en cuenta es el truco de comparar la sucesión que estudiamos con otras más sencillas de analizar. ♦

La aritmética de límites se utiliza, como en el cálculo, para evaluar límites más complejos.

1.2.15 Ejemplo *Evalúese* $\lim\limits_{n\to\infty} \dfrac{n^2 - 3n + 6}{3n^2 + 4n + 2}$.

Solución

$$\begin{aligned}
\lim_{n\to\infty} \frac{n^2 - 3n + 6}{3n^2 + 4n + 2} &= \lim_{n\to\infty} \frac{1 - 3/n + 6/n^2}{3 + 4/n + 2/n^2} \\
&= \frac{\lim\limits_{n\to\infty} 1 - 3 \lim\limits_{n\to\infty} 1/n + 6\left(\lim\limits_{n\to\infty} 1/n\right)^2}{\lim\limits_{n\to\infty} 3 + 4 \lim\limits_{n\to\infty} 1/n + 2\left(\lim\limits_{n\to\infty} 1/n\right)^2} \\
&= \frac{1 - 3\cdot 0 + 6\cdot 0^2}{3 + 4\cdot 0 + 2\cdot 0^2} \\
&= \frac{1}{3} \quad \text{(porque sabemos que } \lim_{n\to\infty} 1/n = 0\text{).}
\end{aligned}$$ ♦

Ahora estudiaremos la construcción de los números reales. El siguiente resultado es básico.

1.2.16 Teorema *Existe un "único" cuerpo ordenado completo al que denominaremos el **sistema de los números reales**.*

El sistema de los números reales se simboliza con $\mathbb{R}$. En este libro $+\infty$ y $-\infty$ *no* están incluidos en $\mathbb{R}$; sin embargo, en otros temas, como en variable compleja y en teoría de integración, es útil agregar $\pm\infty$.

En el teorema anterior, "único" significa que entre todo par de sistemas que cumplan los grupos de axiomas **I–III** y la propiedad de completitud **1.2.9** se puede establecer una correspondencia biunívoca compatible con la suma, la multiplicación y la relación de orden. Ser compatible con la suma, por ejemplo, significa que el número en el segundo sistema correspondiente a la suma de dos números del primero es la suma de los dos números correspondientes en el segundo sistema. Una correspondencia de este tipo se llama ***isomorfismo***.

No expondremos la demostración del teorema **1.2.16** sino que daremos la construcción básica de $\mathbb{R}$ aquí y dejaremos la verificación detallada de las propiedades como un ejercicio. Comenzamos con el cuerpo ordenado de los números racionales $\mathbb{Q}$. Considérese el siguiente conjunto de sucesiones

$$S = \{(x_1, x_2, \ldots) \mid \text{cada } x_n \in \mathbb{Q} \text{ y la sucesión es creciente y está acotada superiormente}\}.$$

Decimos que dos elementos de S son ***equivalentes*** si toda cota superior de una sucesión es cota superior de la otra. El conjunto de todas las sucesiones que son equivalentes a una sucesión dada se denomina ***clase de equivalencia***. Se puede verificar que esto divide al espacio S (o define una partición de éste) en subconjuntos disjuntos. Sea $\mathbb{R}$ el conjunto de todas las clases de equivalencia de S. Los racionales parecen como un subconjunto de $\mathbb{R}$ identificando un número racional r con la clase de todas las sucesiones equivalentes a la sucesión $(r, r, r, \ldots)$. En pocas palabras, los reales son los límites de todas las sucesiones crecientes de racionales que están acotadas superiormente. Ahora debemos definir suma, multiplicación, desigualdad, etcétera, y verificar todos los axiomas. Esta tarea es inmediata, aunque un poco tediosa; como parte de ella uno debe verificar la consistencia, es decir, comprobar que si $x \in \mathbb{R}$ y $(x_1, x_2, \ldots)$ es una sucesión que está en la clase x, entonces esta sucesión realmente converge a x.

La existencia de $\mathbb{R}$ también se puede probar verificando que las expresiones decimales usuales tienen las propiedades requeridas, aunque este método tiene algunas dificultades sutiles. Otos dos métodos clásicos son el geométrico, debido a Richard Dedekind, y el analítico, debido a Georg Cantor. Dedekind elige como elementos básicos las "cortaduras", que son particiones de los racionales en pares de subconjuntos (A, B) con $x < y$

para todo x en A e y en B. Cantor procede como nosotros, estudiando directamente las sucesiones, y construye límites para ellas considerando clases de equivalencia de sucesiones como sus objetos básicos.

Una consecuencia de la propiedad arquimediana de **1.2.11** es que los números racionales son densos en $\mathbb{R}$. Esto significa que dado cualquier número real x, podemos encontrar números racionales tan cercanos a éste como deseemos, lo que confirma nuestra imagen intuitiva de los números densamente distribuidos sobre la recta.

1.2.17 Proposición

$\mathbb{Q}$ es denso en $\mathbb{R}$. Es decir,

i. *Si x e y están en $\mathbb{R}$ y $x < y$, existe $r \in \mathbb{Q}$ tal que $x < r < y$.*

ii. *Si $x \in \mathbb{R}$, $\varepsilon \in \mathbb{R}$ y $\varepsilon > 0$, entonces existe $r \in \mathbb{Q}$ tal que $|x - r| < \varepsilon$.*

Los números racionales son numerables y, por **1.2.17**, hay suficientes como para aproximar cualquier número real. ¿Implica esto que $\mathbb{R}$ es numerable? Aunque parezca sorprendente, la respuesta es no. Éste es otro resultado más de Georg Cantor. Al final del capítulo damos uno de los argumentos más tradicionales para demostrarlo. El argumento se denomina *procedimiento diagonal* y se basa en la representación de los reales en forma de desarrollos decimales infinitos.

1.2.18 Teorema

El intervalo unidad $]0, 1[$ en $\mathbb{R}$ no es numerable.

Un cambio de escala mediante una función lineal $f(x) = a + (b - a)x$ pone el intervalo $]0, 1[$ en correspondencia biunívoca con el intervalo $]a, b[$; por lo tanto, ningún intervalo es numerable. Así pues, aunque **1.2.17** muestra que cualquier intervalo tiene infinitos números racionales, **1.2.18** nos muestra que también debe contener infinitos (un infinito no numerable, de hecho) números irracionales.

1.2.19 Corolario

Si x e y están en $\mathbb{R}$ y $x < y$, entonces el intervalo $]x, y[$ contiene un infinito numerable de números racionales y un infinito no numerable de números irracionales.

Concluimos con una sucesión menos obvia, a la cual recurriremos más adelante como fuente de ejemplos.

1.2.20 Proposición: la sucesión armónica

Sean $x_1 = 1$ y $x_n = 1 + 1/2 + \cdots + 1/n$, $n = 2, 3, \ldots$. Entonces x_n es monótona creciente, pero no está acotada superiormente y, por lo tanto, no converge (escribimos $x_n \to \infty$).

En la sección siguiente estudiaremos la propiedad del supremo. Algunos autores prefieren usarla como la propiedad de completitud. En esta versión, la propiedad de la sucesión monótona aparece como teorema. En nuestro enfoque, que se ha elegido por su carácter intuitivo, ¡es la propiedad del supremo la que aparece como teorema!

Ejercicios de §1.2

1. En el ejemplo **1.2.10**, sea $\lambda = \text{lím}_{n\to\infty} x_n$.

a. Demuéstrese que λ es una raíz de $\lambda^2 - \lambda - 2 = 0$.

b. Encuéntrese $\text{lím}_{n\to\infty} x_n$.

2. Pruébese que $3^n/n!$ converge a 0.

3. Sea $x_n = \sqrt{n^2+1} - n$. Calcúlese $\text{lím}_{n\to\infty} x_n$.

4. Sea x_n una sucesión monótona creciente tal que $x_{n+1} - x_n \leq 1/n$. ¿Debe x_n ser convergente?

5. Sea **F** un cuerpo ordenado en el cual toda sucesión monótona estrictamente creciente acotada superiormente converge. Pruébese que $\mathbb{F}$ es completo.

§1.3 Supremos

El axioma de completitud se puede enunciar de muchas otras formas equivalentes, pero para ello es necesario ampliar la terminología.

1.3.1 Definiciones *Sea $S \subset \mathbb{R}$. Un número b es una **cota superior** de S si para todo $x \in S$ se cumple $x \leq b$.*

*Un número b es **supremo** de S si, en primer lugar, b es una cota superior de S y, en segundo lugar, b es menor o igual que cualquier otra cota superior de S (véase la figura 1.3-1). El supremo de S (también llamado **cota superior mínima** de S) se denota* sup *S o* sup(*S*). *Si $S \subset \mathbb{R}$ no está acotado superiormente (no tiene cotas superiores) o S es vacío, decimos que* sup *S es **infinito** y escribimos* $\sup S = +\infty$.[1]

[1] Una interpretación literal de la definición nos puede tentar a definir $\sup(\varnothing) = -\infty$. Sin embargo, guiados por las propiedades de la proposición **1.3.3**, que se enuncia más adelante, hemos preferido definir $\sup(\varnothing) = +\infty$.

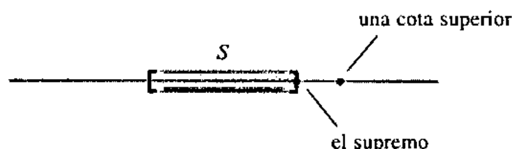


FIGURA 1.3-1 Supremo

El conjunto $]a, b[= \{x \in \mathbb{R} \mid a < x < b\}$ se denomina ***intervalo abierto***, mientras que el conjunto $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$ se denomina ***intervalo cerrado***. Por ejemplo, el intervalo cerrado $[0,1]$, el intervalo abierto $]0, 1[$ y el conjunto de todos los racionales menores que 1 tienen como supremo 1. Éstos son conjuntos relativamente sencillos. Conviene tener presente que la teoría que desarrollamos en esta sección también se aplica a conjuntos complicados, como el de los números en $[0, 1]$ con expresión decimal $0.a_1a_2a_3...$, donde a_i es 5, 6 o 7 para cada i, un conjunto que no se puede visualizar con facilidad.

> **Nota.** En este libro los intervalos abiertos se denotan con $]a, b[$ en lugar de (a, b); esta notación europea evita confundirlos con los pares ordenados.

Un conjunto *S sólo puede tener un supremo cuando mucho*. En efecto: si b y b' son supremos, como b es menor o igual que cualquier otra cota superior, $b \leq b'$; de manera similar $b' \leq b$, de donde podemos concluir que $b = b'$. Por esta razón podemos hablar de *el* supremo.

No es necesario que el supremo sea un elemento del conjunto. Por ejemplo, el supremo de $S =]0, 1[$, que es 1, no pertenece a S.

Un conjunto *no tiene por qué tener supremo*. Por ejemplo, el sistema de los números reales no tiene supremo, y tampoco lo tiene el conjunto de los enteros positivos. En el caso "degenerado" del conjunto vacío consideramos que cualquier número es una cota superior.

Si b es una cota superior del conjunto S y b ∈ S, b es el supremo. Para ver esto, obsérvese que si d es cualquier cota superior de S, entonces $b \in S$ implica $b \leq d$, como se requiere.

Una alternativa útil a la definición de supremo se enuncia a continuación.

1.3.2 Proposición *Sea $S \subset \mathbb{R}$ no vacío. Entonces $b \in \mathbb{R}$ es el supremo de S sii b es una cota superior y para todo $\varepsilon > 0$ existe $x \in S$ tal que $x > b - \varepsilon$.*

La prueba se puede encontrar al final del capítulo; sin embargo, el teorema debe ser "obvio", porque b está justo en la "cima" (es decir, a la "derecha") del conjunto S y no

hay "huecos" entre b y el conjunto S, por tanto para cualquier $\varepsilon > 0$ podemos tomar x justo debajo de b con una separación menor que ε. [¡Cuidado!, este tipo de argumento de plausibilidad tiene el propósito de hacer entender el enunciado y no debe confundirse con una demostración rigurosa.]

De manera similar, una ***cota inferior*** para un conjunto S es un número b tal que $b \leq x$ para todo $x \in S$. Más aún, b es ***ínfimo*** sii es una cota inferior y para cualquier cota inferior c de S, $c \leq b$. Como ocurre con el supremo, el ínfimo es único, cuando existe. El ínfimo también se llama ***cota inferior máxima*** y se denota inf S o inf(S). Como en **1.3.2**, un número c es el ínfimo de un conjunto S sii c es una cota inferior y para todo $\varepsilon > 0$ existe un $x \in S$ tal que $x < c + \varepsilon$. Si $S \subset \mathbb{R}$ no está acotado inferiormente o es vacío, escribimos $\inf(S) = -\infty$.

Hay algunas propiedades más o menos evidentes enunciadas en la siguiente proposición.

1.3.3 Proposición *Supongamos que $A \subset B \subset \mathbb{R}$. Entonces*

$$\inf B \leq \inf A \leq \sup A \leq \sup B.$$

En **1.3.3** suponemos que el ínfimo y el supremo existen, o usamos las convenciones $-\infty < x < +\infty$ para cualquier x en $\mathbb{R}$. El problema de la existencia, que es más sutil, se trata en el teorema siguiente.

1.3.4 Teorema *Las siguientes afirmaciones son válidas en $\mathbb{R}$:*

i. ***Propiedad del supremo*** *Sea S un conjunto no vacío de $\mathbb{R}$ que tiene una cota superior; entonces S tiene supremo en $\mathbb{R}$.*

ii. ***Propiedad del ínfimo*** *Sea P un conjunto no vacío de $\mathbb{R}$ que tiene una cota inferior; entonces P tiene ínfimo en $\mathbb{R}$.*

La demostración se da al final del capítulo; sin embargo, este resultado también debe ser evidente. En efecto: si un subconjunto acotado de $\mathbb{R}$ no tuviera supremo, habría un "hueco" en lo alto del conjunto y una sucesión creciente de elementos de S no convergería a un elemento de $\mathbb{R}$. De manera similar obtenemos la propiedad **ii**. Usando los métodos de la demostración, no es difícil probar que las condiciones **i** y **ii** son equivalentes al axioma de completitud para un cuerpo ordenado.

1.3.5 Ejemplo *Considérese el conjunto $S = \{x \in \mathbb{R} \mid x^2 + x < 3\}$. Encuéntrense* sup S e inf S.

Solución Consideremos la gráfica de $y = x^2 + x$ (figura 1.3-2). Utilizando el cálculo elemental vemos que para $x = -1/2$, y alcanza un mínimo. Entonces S se puede dibujar como en la figura 1.3-2.

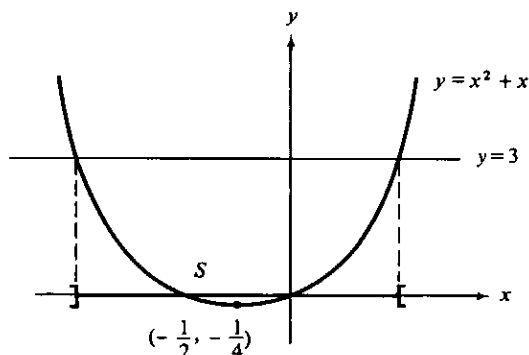


FIGURA 1.3-2 El conjunto S del ejemplo **1.3.5**

Claramente, el supremo y el ínfimo se alcanzan cuando $x^2 + x = 3$ o, por la fórmula cuadrática, cuando

$$x = \frac{-1 \pm \sqrt{1+12}}{2} = \frac{-1 \pm \sqrt{13}}{2}.$$

Por consiguiente,

$$\sup S = \frac{\sqrt{13}-1}{2} \quad \text{e} \quad \inf S = \frac{-(\sqrt{13}+1)}{2}. \quad ♦$$

Ejercicios de §1.3

1. Sea $S = \{x \mid x^3 < 1\}$. Encuéntrese sup S. ¿Está S acotado inferiormente?
2. Considérese una sucesión creciente x_n que está acotada superiormente y que converge a x. Sea $S = \{x_n \mid n = 1, 2, 3, \ldots\}$. Proporciónese un argumento de plausibilidad de que $x = \sup S$.
3. Si $P \subset Q \subset \mathbb{R}$, $P \neq \varnothing$, y P y Q están acotados superiormente, pruébese que sup $P \leq$ sup Q.
4. Sean $A \subset \mathbb{R}$ y $B \subset \mathbb{R}$ acotados inferiormente y definamos $A + B = \{x + y \mid x \in A$ e $y \in B\}$. ¿Es cierto que $\inf(A + B) = \inf A + \inf B$?
5. Sea $S \subset [0, 1]$ el conjunto que consta de todas las expresiones decimales infinitas $x = 0.a_1a_2a_3\ldots$, donde todos los dígitos excepto un número finito de ellos son 5 o 6. Encuéntrese sup S.

§1.4 Sucesiones de Cauchy

Aunque todas las sucesiones monótonas y acotadas convergen, no todas las sucesiones que convergen son monótonas. El propósito de esta sección es caracterizar todas las sucesiones que convergen. Los términos de una sucesión que converge a x se acercan a x, y, por lo tanto, también deben acercarse entre sí. Esta propiedad recibió su nombre de Augustin Louis Cauchy (1789–1857), cuyas ideas forman una buena parte de la base teórica del cálculo y el análisis moderno. Muchas de esas ideas están contenidas en su trabajo de 1823 *Résumé des leçons donées à l'école royale polytechnique sur le calcul infinitésimal*. (Véase *The Origins of Cauchy's Rigorous Calculus* de J. Grabiner, MIT Press, Cambridge, Mass., 1981.)

1.4.1 Definición *Una sucesión x_n de números reales es una* ***sucesión de Cauchy*** *si para todo número $\varepsilon > 0$ existe un entero N (que depende de ε) tal que $|x_n - x_m| < \varepsilon$ siempre que $n \geq N$ y $m \geq N$.*

El significado intuitivo de esta condición es que la sucesión se "amontona", es decir, que todos los términos de la sucesión están tan cerca como queramos, unos de otros, para índices suficientemente grandes.

Si x_n converge a x, podemos afirmar que x_n es una sucesión de Cauchy. En efecto, dado $\varepsilon > 0$, elegimos N tal que $|x_n - x| < \varepsilon/2$ si $n \geq N$. Si $n,m \geq N$, entonces $|x_n - x_m| = |x_n - x + x - x_m| \leq |x_n - x| + |x - x_m| < \varepsilon/2 + \varepsilon/2 = \varepsilon$, lo cual prueba nuestra afirmación. En esta demostración hemos usado la desigualdad triangular $|y - z| \leq |y| + |z|$. Resumiendo:

1.4.2 Proposición *Toda sucesión convergente es una sucesión de Cauchy.*

Los dos teoremas siguientes son centrales para la mayor parte del análisis. El primero, que es la base de muchas de las ideas sobre la compacidad del capítulo 3, se aprovecha de la propiedad de que, aunque una sucesión no converja, algunas de sus partes tal vez lo hagan. Consideremos la sucesión $1, -1, 1, -1, 1, -1, \ldots$ definida para $n = 0, 1, 2, 3, \ldots$ por $x_n = (-1)^n$. Esta sucesión no converge, pero si nos fijamos sólo en los términos con índice par: $x_0, x_2, x_4, \ldots$ tenemos la sucesión constante $1, 1, 1, \ldots$, que sí converge. Si en su lugar nos fijamos sólo en los términos con índice impar: $x_1, x_3, x_5, \ldots$ tenemos otra sucesión constante $-1, -1, -1, \ldots$, que también converge. Decimos que hemos encontrado *subsucesiones* convergentes. Una ***subsucesión*** de una sucesión $x_0, x_1, x_2, \ldots$ es una sucesión que se construye eliminando parte de los términos de la sucesión de partida, mientras los demás se conservan *sin alterar su orden*. De manera más precisa: seleccionamos los índices $n(1) < n(2) < n(3) < \cdots$ y construimos la sucesión con estos índices:

$$x_{n(1)}, x_{n(2)}, x_{n(3)}, \ldots$$

El teorema **1.4.3** es nuestro primer encuentro con una observación muy importante conocida como la ***propiedad de Bolzano-Weierstrass***: si una sucesión en $\mathbb{R}$ está acotada, se debe acumular en alguna parte y tener una subsucesión (posiblemente muchas) que converja a algún punto de $\mathbb{R}$. Este hecho depende crucialmente de la completitud, ya que el posible límite ha de estar presente. El segundo teorema está muy relacionado con el primero y nos da la versión de la completitud que resulta ser más apropiada para las generalizaciones.

1.4.3 Teorema *Toda sucesión acotada en $\mathbb{R}$ tiene una subsucesión que converge a algún punto de $\mathbb{R}$.*

1.4.4 Teorema *Toda sucesión de Cauchy en $\mathbb{R}$ converge a un elemento de $\mathbb{R}$.*

1.4.5 Corolario *Si a y b están en $\mathbb{R}$ y $a < b$, entonces toda sucesión de puntos en el intervalo $[a, b] = \{x \mid a \leq x \leq b\}$ tiene una subsucesión que converge a algún punto de $[a, b]$.*

La propiedad expresada en el teorema **1.4.4** se llama ***completitud en el sentido de Cauchy***. En ocasiones se da como definición de completitud, pero no puede, por sí sola, sustituir al axioma del supremo, a menos que se adopte junto con la propiedad arquimediana. Un cuerpo arquimediano completo en el sentido de Cauchy satisface la propiedad del supremo. (Sin la propiedad arquimediana, podría no satisfacerla. Una sucesión decreciente y acotada como $1, 1/2, 1/3, 1/4, 1/5, \ldots$ podría no converger a nada. El único candidato posible al límite es 0, pero sin la propiedad arquimediana no tiene necesariamente que converger a 0. Por un motivo similar tampoco sería una sucesión de Cauchy.)

Una idea intuitiva del teorema **1.4.4** es la siguiente. Si ignoramos los primeros N términos de una sucesión de Cauchy, sabemos que los términos restantes se amontonan. Conforme desechemos más y más términos, el resto de la sucesión se aprieta más y se agolpa contra algún número límite, el límite de la sucesión. Ver con más precisión cómo ocurre esto requiere un mayor cuidado y, por lo tanto, nuestro único recurso es la demostración. Los siguientes lemas contienen la estrategia para probar el teorema **1.4.4** a partir del **1.4.3**.

1.4.6 Lema *Toda sucesión de Cauchy está acotada.*

1.4.7 Lema *Si una subsucesión de una sucesión de Cauchy converge a x, entonces la sucesión misma converge a x.*

El lema **1.4.6** es creíble, porque una sucesión de Cauchy tiene términos que se amontonan (digamos, a menos de una distancia 1 de algún punto), y dejar fuera un número finito de puntos no altera el hecho de ser acotada. La idea del lema **1.4.7** es que una subsucesión convergente debe arrastrar a la sucesión completa hacia el límite, ya que todos los términos están muy cerca uno del otro para valores grandes del índice.

Para probar **1.4.4** a partir de **1.4.3**, obsérvese que, según **1.4.6**, toda sucesión de Cauchy está acotada y, por lo tanto, de **1.4.3** se sigue que tiene una subsucesión convergente. Por **1.4.7**, toda la sucesión de Cauchy ha de ser convergente.

En cuanto al teorema **1.4.3**, es plausible que una sucesión acotada se tenga que amontonar en alguna parte y, gracias a este hecho, podemos extraer de ella una subsucesión convergente. El principal problema técnico que enfrentamos en la demostración rigurosa es cómo usar la completitud de $\mathbb{R}$; para ser más precisos, cómo relacionar la sucesión dada con una monótona. Trate el lector de pensar en algunas posibilidades antes de leer la demostración (dichos intentos, aun si no tienen éxito, pueden ayudar a aguzar las habilidades matemáticas).

1.4.8 Ejemplo *Sea x_n una sucesión de números reales con la propiedad de que $|x_n - x_{n+1}| \leq 1/2^n$. Demuéstrese que x_n converge.*

Solución Si demostramos que x_n es una sucesión de Cauchy, el resultado será consecuencia del teorema **1.4.4**. Podemos escribir, por la desigualdad triangular,

$$\begin{aligned}|x_n - x_{n+k}| &\leq |x_n - x_{n+1}| + |x_{n+1} - x_{n+2}| + \cdots + |x_{n+k-1} - x_{n+k}| \\ &\leq \frac{1}{2^n} + \frac{1}{2^{n+1}} + \cdots + \frac{1}{2^{n+k-1}} \leq \frac{2}{2^n},\end{aligned}$$

puesto que $a + ar + ar^2 + \cdots + ar^n = a(1 - r^{n+1})/(1 - r) < a/(1 - r)$ si $0 < r < 1$. Entonces $|x_n - x_m| \leq 1/2^{n-1}$ si $m \geq n$. Dado $\varepsilon > 0$, si elegimos N de tal manera que $1/(2^{N-1}) < \varepsilon$, entonces x_n satisface la definición de una sucesión de Cauchy. ♦

Ejercicios de §1.4

1. Supongamos que x_n satisface $|x_n - x_{n+1}| < 1/3^n$. Pruébese que x_n converge.
2. Demuéstrese que la sucesión $x_n = e^{\operatorname{sen}(5n)}$ tiene una subsucesión convergente.
3. Encuéntrese una sucesión acotada con tres subsucesiones que converjan a tres números diferentes.

4. Sea x_n una sucesión de Cauchy. Supongamos que para todo $\varepsilon > 0$ existe algún $n > 1/\varepsilon$ tal que $|x_n| < \varepsilon$. Pruébese que $x_n \to 0$.

5. ¿Cierto o falso?: si x_n es una sucesión de Cauchy, entonces para n y m suficientemente grandes, $d(x_{n+1}, x_{m+1}) \leq d(x_n, x_m)$.

§1.5 Puntos límite; lím inf y lím sup

Una herramienta útil en el estudio de las sucesiones convergentes es la noción de punto límite.

1.5.1 Definición *Un punto x se llama* ***punto límite*** *de la sucesión x_n si para toda $\varepsilon > 0$ hay infinitos valores de n tales que $|x_n - x| < \varepsilon$.*

Por ejemplo, 1 y -1 son puntos límite de la sucesión $1, -1, 1, -1, 1, -1, \ldots$; obsérvese que esta sucesión *no* converge. No obstante, la proposición siguiente muestra que hay una relación entre convergencia y punto límite.

1.5.2 Proposición *Sea x_n una sucesión en $\mathbb{R}$ y sea $x \in \mathbb{R}$.*

i. *x es un punto límite de x_n si y sólo si para todo $\varepsilon > 0$ y para todo N, existe un índice $n > N$ tal que $|x_n - x| < \varepsilon$.*

ii. *x es un punto límite de x_n si y sólo si existe una subsucesión de x_n que converja a x.*

iii. *$x_n \to x$ si y sólo si toda subsucesión de x_n converge a x.*

iv. *$x_n \to x$ si y sólo si la sucesión está acotada y x es su único punto límite.*

v. *$x_n \to x$ si y sólo si toda subsucesión de x_n tiene a su vez una subsucesión que converge a x.*

Considérese la sucesión $1, 0, -1, 1, 0, -1, 1, 0, -1, \ldots$. Esta sucesión no converge. Tiene tres puntos límite 1, 0 y -1. De ellos, 1 y -1 son los más interesantes, por ser el más grande y el más pequeño de los puntos límite. Parece obvio a primera vista, aunque

menos obvio si se piensa un poco, que toda sucesión acotada de números reales tenga un punto límite que es el mayor y otro que es el menor. Vamos a intentar introducir este hecho en una definición para sucesiones generales en $\mathbb{R}$.

1.5.3 Definición *Sea x_n una sucesión acotada superiormente en $\mathbb{R}$. El **límite superior** de x_n es el mayor punto límite de x_n, es decir, el supremo de los puntos límite. Si el conjunto de puntos límite es vacío, definimos* lím sup $x_n = -\infty$; *si la sucesión no está acotada superiormente, definimos* lím sup $x_n = +\infty$. *De manera similar, si x_n está acotada inferiormente, el **límite inferior** de x_n es el menor punto límite de x_n, es decir, el ínfimo de los puntos límite. Si el conjunto de puntos límite es vacío, definimos* lím inf $x_n = +\infty$; y *si la sucesión no está acotada inferiormente, definimos* lím inf $x_n = -\infty$.

El límite superior se denota como sigue:

$$\text{lím sup}\, x_n, \quad \text{lím sup}_{n\to\infty} x_n, \quad \text{o} \quad \overline{\text{lím}}\; x_n$$

y el límite inferior

$$\text{lím inf}\, x_n, \quad \text{lím inf}_{n\to\infty} x_n, \quad \text{o} \quad \underline{\text{lím}}\; x_n.$$

1.5.4 Ejemplos

a. Para la sucesión 1, 0, −1, 1, 0, −1, 1, 0, −1, . . . , lím inf $x_n = -1$ y lím sup $x_n = 1$.

b. La sucesión $x_n = 1/n$ converge a 0, que es el único punto límite. Por tanto, lím x_n = lím sup x_n = lím inf $x_n = 0$.

c. Sea $x_n = 0$ si n es par y $x_n = n$ si n es impar; es decir, $x_n = 0, 1, 0, 3, 0, 5, 0, 7, \ldots$. En este caso, lím inf $x_n = 0$. La sucesión no está acotada superiormente y, por lo tanto, escribimos lím sup $x_n = +\infty$.

d. Sea $x_n = n$, es decir, 0, 1, 2, 3, 4, ... Aquí, lím sup x_n = lím inf $x_n = +\infty$.

e. Aun si una sucesión está acotada, el lím inf no es, necesariamente, el ínfimo del conjunto $\{x_1, x_2, x_3, \ldots\}$, ni el lím sup su supremo. Esto se debe a que estamos interesados sólo en los puntos límite, y podemos eliminar cualquier número finito de términos sin alterar el conjunto de puntos límite.

Para un ejemplo más complicado, considérese la sucesión x_n definida por

$$\begin{aligned} 1 + 1/j &\quad \text{si } n = 5j \\ 1 - 1/j &\quad \text{si } n = 5j + 1 \\ 0 &\quad \text{si } n = 5j + 2 \\ -1 + 1/j &\quad \text{si } n = 5j + 3 \\ -1 - 1/j &\quad \text{si } n = 5j + 4 \end{aligned}$$

donde $j = 2, 3, 4, 5, 6, 7, 8, \ldots$ y $n = 10, 11, 12, 13, \ldots$. La sucesión comienza con

$$\frac{3}{2}, \frac{1}{2}, 0, -\frac{1}{2}, -\frac{3}{2}, \frac{4}{3}, \frac{2}{3}, 0, -\frac{2}{3}, -\frac{4}{3}, \frac{5}{4}, \frac{3}{4}, 0, -\frac{3}{4}, -\frac{5}{4}, \ldots,$$

como se muestra en la figura 1.5-1.

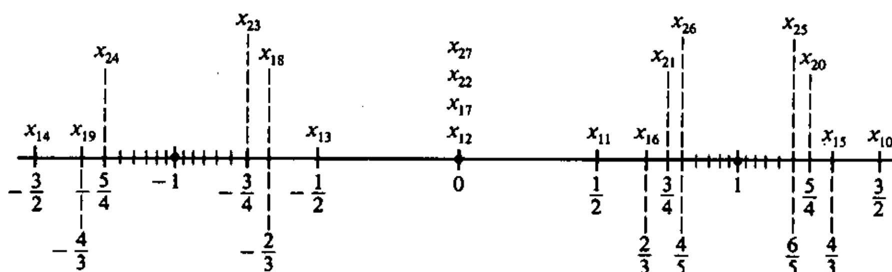


FIGURA 1.5-1 La sucesión del ejemplo **1.5.4**, apartado e

Hay subsucesiones que convergen a -1, 0 y 1, y subsucesiones que se aproximan a $+1$ y -1 por arriba y por abajo. Entonces, $\text{lím sup}\, x_n = 1$ y $\text{lím inf}\, x_n = -1$.

Un ejemplo un poco más sencillo es:

f. $x_n = (-1)^n(1 + 1/n)$, $n = 1, 2, 3, 4, \ldots$. Los primeros términos son

$$-2, \frac{3}{2}, -\frac{4}{3}, \frac{5}{4}, -\frac{6}{5}, \frac{7}{6}, -\frac{8}{7}, \frac{9}{8}, \ldots.$$

Los términos con índice impar crecen hacia -1 y los términos con índice par decrecen hacia 1, como se observa en la figura 1.5-2. Entonces, $\text{lím inf}\, x_n = -1$ y $\text{lím sup}\, x_n = 1$. ♦

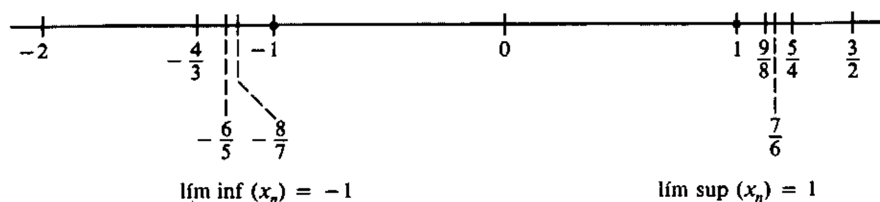


FIGURA 1.5-2 La sucesión del ejemplo **1.5.4**, apartado **f**

En lo que sigue téngase presente lo siguiente: la idea clave es que un punto límite de una sucesión sigue siendo un punto límite aun si se elimina un número finito de puntos. Si x es un punto límite de $x_1, x_2, x_3, x_4, \ldots$, entonces también es un punto límite de $x_{n+1}, x_{n+2}, x_{n+3}, x_{n+4}, \ldots$. Al usar lím inf y lím sup, en ocasiones es útil tener una caracterización directa del tipo ε-N.

1.5.5 Proposición *Sea x_n una sucesión en $\mathbb{R}$. Entonces*

i. *Si x_n está acotada inferiormente, un número a es igual al* lím inf x_n *si y sólo si*

a. *Para todo $\varepsilon > 0$ existe un N tal que $a - \varepsilon < x_n$ siempre que $n \geq N$, y*

b. *Para todo $\varepsilon > 0$ y para todo M existe un $n > M$ tal que $x_n < a + \varepsilon$.*

ii. *Si x_n está acotada superiormente, un número b es igual al* lím sup x_n *si y sólo si*

a. *Para todo $\varepsilon > 0$ existe un N tal que $x_n < b + \varepsilon$ siempre que $n \geq N$, y*

b. *Para todo $\varepsilon > 0$ y para todo M existe un $n > M$ tal que $b - \varepsilon < x_n$.*

Un examen atento de la figura 1.5-2 debe hacer esta proposición plausible. Usándola podemos desarrollar otra caracterización del lím inf y el lím sup. Dada una sucesión $x_1, x_2, x_3, \ldots$ en $\mathbb{R}$, sean $S_n = \{x_{n+1}, x_{n+2}, x_{n+3}, \ldots\}$, $a_n = \inf S_n$ y $b_n = \sup S_n$ (recordemos que si S_n no está acotado superiormente, entonces $\sup S_n = +\infty$ y si S_n no está acotado inferiormente, entonces $\inf S_n = -\infty$). Entonces, $S_1 \supset S_2 \supset S_3 \supset S_4 \supset S_5 \ldots$ Si $n \leq k$, $S_k \subset S_n$, de donde $a_k \leq b_k \leq b_n$. Si $k \leq n$, $S_n \subset S_k$ así que $a_k \leq a_n \leq b_n$. Por consiguiente, las a_n y las b_n están ordenadas de la siguiente forma

$$a_1 \leq a_2 \leq a_3 \leq \cdots \leq b_3 \leq b_2 \leq b_1.$$

(Obsérvese que $+\infty$ y $-\infty$ son valores legítimos para a_n y b_n.) El lector debería calcular algunos de estos valores para la sucesión del ejemplo **1.5.4,** apartado **f**.

1.5.6 Proposición *Sea x_n una sucesión dada en $\mathbb{R}$. Entonces*

$$\text{lím sup}\, x_n = \inf_n\{\sup S_n\} = \inf\{b_1, b_2, b_3, b_4, \ldots\}$$
$$\text{lím inf}\, x_n = \sup_n\{\inf S_n\} = \sup\{a_1, a_2, a_3, a_4, \ldots\}.$$

Si $S_n = \{x_{n+1}, x_{n+2}, x_{n+3}, \ldots\}$, $a_n = \inf S_n$ y $b_n = \sup S_n$, podemos reescribir esto como

$$\text{lím sup}\, x_n = \inf\{\sup\{x_{n+1}, x_{n+2}, x_{n+3}, \ldots\} \mid n = 1, 2, 3, \ldots\}$$
$$\text{lím inf}\, x_n = \sup\{\inf\{x_{n+1}, x_{n+2}, x_{n+3}, \ldots\} \mid n = 1, 2, 3, \ldots\}.$$

A continuación damos algunas propiedades generales.

1.5.7 Proposición *Sea x_n una sucesión dada en $\mathbb{R}$. Entonces*

i. lím inf $x_n \le$ lím sup x_n.

ii. *Si $x_n \le M$ para todo n, entonces* lím sup $x_n \le M$.

iii. *Si $M \le x_n$ para todo n, entonces* lím inf $x_n \ge M$.

iv. lím sup $x_n = +\infty$ *si y sólo si x_n no está acotada superiormente.*

v. lím inf $x_n = -\infty$ *si y sólo si x_n no está acotada inferiormente.*

vi. *Si x es un punto límite de x_n, entonces* lím inf $x_n \le x \le$ lím sup x_n.

vii. *Si $a =$* lím inf x_n *es finito, entonces es un punto límite de x_n.*

viii. *Si $b =$* lím sup x_n *es finito, entonces es un punto límite de x_n.*

ix. $x_n \to x \in \mathbb{R}$ *si y sólo si* lím sup $x_n =$ lím inf $x_n = x \in \mathbb{R}$.

Ejercicios de §1.5

1. Sea $x_n = 3 + (-1)^n(1 + 1/n)$. Calcúlense lím inf x_n y lím sup x_n.

2. Encuéntrese una sucesión x_n tal que lím sup $x_n = 5$ y lím inf $x_n = -3$.

3. Sea x_n una sucesión tal que lím sup $x_n = b \in \mathbb{R}$ y lím inf $x_n = a \in \mathbb{R}$. Demuéstrese que x_n tiene subsucesiones u_n y l_n tales que $u_n \to b$ y $l_n \to a$.

4. Supongamos que lím sup $x_n = 2$. ¿Verdadero o falso?: si n es suficientemente grande $x_n > 1.99$.

5. ¿Verdadero o falso?: si lím sup $x_n = b$, entonces para n suficientemente grande, $x_n \le b$.

§1.6 El espacio euclídeo

En este libro trabajamos sobre todo con espacios euclídeos uni, bi o tridimensionales. En muchas aplicaciones y al desarrollar el análisis en el espacio tridimensional, aparecen, de manera natural, espacios de dimensión mayor. En consecuencia, es importante estudiar el caso general, pero regresaremos a los casos de dimensión uno, dos o tres para visualizar las propiedades y adquirir intuición. Comencemos con la definición formal.

1.6.1 Definición *El espacio euclídeo n-dimensional, denotado* $\mathbb{R}^n$, *consta de todas las n-uplas ordenadas de números reales. Simbólicamente,*

$$\mathbb{R}^n = \{(x_1, \ldots, x_n) \mid x_1, \ldots, x_n \in \mathbb{R}\}.$$

Por lo tanto, $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$ *(n veces) es el producto cartesiano de* $\mathbb{R}$ *consigo mismo n veces.*

Los elementos de $\mathbb{R}^n$, en general, se denotan con una sola letra que representa una n-upla, como $x = (x_1, x_2, \ldots, x_n)$, y decimos que x es un ***punto*** de $\mathbb{R}^n$.

La suma y el producto por un escalar de n-uplas se definen mediante

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

y

$$a(x_1, \ldots, x_n) = (ax_1, \ldots, ax_n) \text{ para } a \in \mathbb{R}.$$

El significado geométrico de estas operaciones se ve en la figura 1.6-1 para $n = 3$.

Los espacios $\mathbb{R}^2$ y $\mathbb{R}^3$ y, en general, $\mathbb{R}^n$, son ejemplos de una estructura del álgebra lineal llamada ***espacio vectorial***. De la misma manera que lo hicimos con los cuerpos, identificamos las propiedades esenciales de los espacios conocidos, $\mathbb{R}^2$ y $\mathbb{R}^3$, y las volcamos en una definición. Básicamente, un espacio vectorial es una colección de objetos, llamados ***vectores***, que se pueden sumar, restar y multiplicar por números.

1.6.2 Definición *Un **espacio vectorial real** V es un conjunto de elementos llamados **vectores**, dotado de las operaciones de **suma de vectores*** $+ : V \times V \to V$ *y de **multiplicación por un escalar*** $\cdot : \mathbb{R} \times V \to V$ *que cumplen las siguientes propiedades:*

i. $v + w = w + v$ *para todo* v *y* w *de* V. ***conmutatividad***

ii. $(v + u) + w = v + (u + w)$. ***asociatividad***

iii. *Existe un vector cero,* 0, *tal que* $v + 0 = v$ *para todo* v *de* V. ***vector cero***

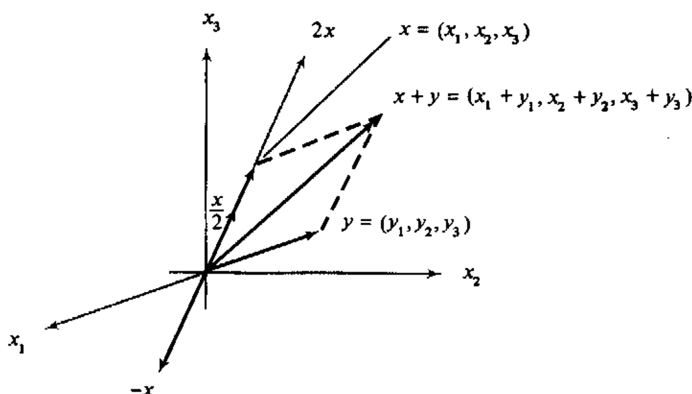


FIGURA 1.6-1 Suma y multiplicación por un escalar

iv. *Para cada v de V existe un vector $-v$ tal que $v + (-v) = 0$.* ***negativos***

v. *Para cada $\lambda \in \mathbb{R}$ y v, w de V, $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$.* ***distributividad***

vi. *Para λ, $m \in \mathbb{R}$ y v de V, $\lambda(m \cdot v) = (\lambda m) \cdot v$.* ***asociatividad***

vii. *Para λ, $m \in \mathbb{R}$ y v de V, $(\lambda + m) \cdot v = \lambda \cdot v + m \cdot v$.* ***distributividad***

viii. $1 \cdot v = v$ *para todo $v \in V$.* ***identidad multiplicativa***

Un espacio vectorial sobre un cuerpo general $\mathbf{F}$ se define de la misma manera, pero sustituyendo $\mathbf{F}$ por $\mathbb{R}$ en todas partes. En análisis nos ocupamos principalmente de espacios vectoriales sobre $\mathbb{R}$, que se denominan ***espacios vectoriales reales***, y sobre los complejos, C, a los que llamaremos ***espacios vectoriales complejos*** (véase §**1.8** para estudiar C). Un subconjunto de V se llama ***subespacio*** (o ***subespacio lineal*** o ***subespacio vectorial***) si él mismo es un espacio vectorial con las mismas operaciones.

1.6.3 Lema *Si W es un subconjunto de un espacio vectorial V sobre un cuerpo* $\mathbf{F}$, *entonces W es un subespacio vectorial sobre* $\mathbf{F}$ *si y sólo si $\lambda v + \mu w \in W$ siempre que λ y μ estén en F y v y w estén en W.*

El lector puede intentar la demostración o repasar la sección correspondiente en cualquier libro de álgebra lineal.

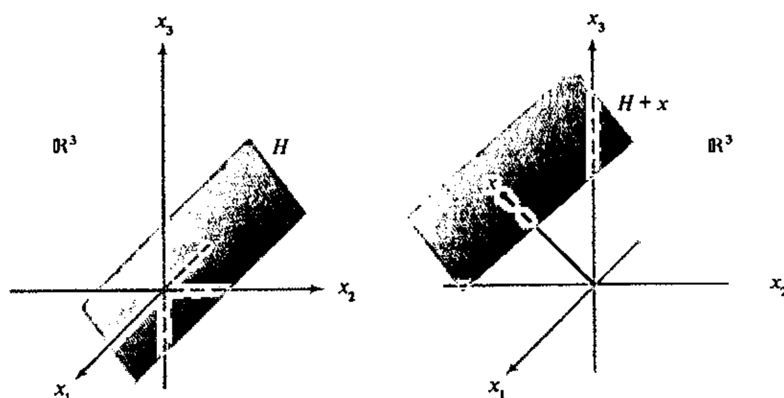


FIGURA 1.6-2 Hiperplano e hiperplano afín

En particular, un subespacio lineal $(n-1)$-dimensional de $\mathbb{R}^n$ se llama ***hiperplano***. Un ***hiperplano afín*** es un conjunto de la forma $x + H$, donde H es un hiperplano, $x \in \mathbb{R}^n$ y $x + H$ significa el conjunto de todos los elementos de la forma $x + y$ donde y pertenece a H, es decir, $x + H = \{x + y \mid y \in H\}$. Véase la figura 1.6-2.

1.6.4 Teorema *El espacio euclídeo n-dimensional, con las operaciones suma y multiplicación por un escalar anteriormente definidas, es un espacio vectorial de dimensión n.*

La demostración es una verificación directa de los axiomas de espacio vectorial, que pospondremos hasta el ejercicio 16 al final del capítulo. Este teorema no debería sorprender; después de todo, un espacio vectorial es una abstracción de las propiedades básicas de los vectores en un espacio euclídeo. Podemos demostrar que $\mathbb{R}^n$ tiene dimensión n mostrando una ***base*** con n vectores; por ejemplo, la ***base canónica*** $\{e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 0, 1)\}$.

Usando la base canónica, las componentes de $x = (x_1, \ldots, x_n)$ son justamente $x_1, \ldots, x_n$. Usando otra base de $\mathbb{R}^n$, las componentes serían diferentes. Esto significa que, si $e_1, \ldots, e_n$ denota la base canónica, $x = \sum_{i=1}^{n} x_i e_i$, pero si $f_1, \ldots, f_n$ es otra base, $x = \sum_{i=1}^{n} y_i f_i$ para ciertos números $y_1, \ldots, y_n$ (probablemente diferentes). No hemos definido los términos "base" o "componente" en general. Para ello remitimos al lector a su curso de álgebra lineal.

1.6.5 Definición *La **longitud** o **norma** de un vector x de $\mathbb{R}^n$ se define mediante*

$$\|x\| = \left\{\sum_{i=1}^{n} x_i^2\right\}^{1/2}.$$

donde $x = (x_1, \ldots, x_n)$. *La* ***distancia*** *entre dos vectores x e y es el número real*

$$d(x, y) = \|x - y\| = \left\{ \sum_{i=1}^{n} (x_i - y_i)^2 \right\}^{1/2}.$$

El producto interno *de x e y es el número real definido por*

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i.$$

En consecuencia tenemos $\|x\|^2 = \langle x, x \rangle$. En $\mathbb{R}^3$, el lector conoce otra expresión para $\langle x, y \rangle$, a saber, $\langle x, y \rangle = \|x\| \, \|y\| \cos\theta$, donde $\cos\theta$ es el coseno del ángulo formado por x e y. Véase la figura 1.6-3.

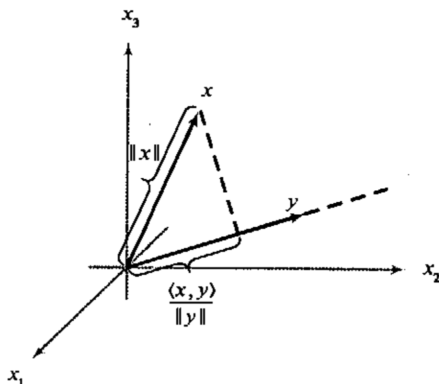


FIGURA 1.6-3 Longitud y producto interno

A continuación resumimos las propiedades básicas de estas operaciones:

1.6.6 Teorema *Para vectores de* $\mathbb{R}^n$*, tenemos*

I. *Propiedades del producto interno:*

i. $\langle x, x \rangle \geq 0$. ***positividad***

ii. $\langle x, x \rangle = 0$ *sii* $x = 0$. ***no degeneración***

iii. $\langle x, y + w \rangle = \langle x, y \rangle + \langle x, w \rangle$. ***distributividad***

iv. $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$ *para todo* $\alpha \in \mathbb{R}$. ***multiplicatividad***

v. $\langle x, y\rangle = \langle y, x\rangle$. ***simetría***

II. *Propiedades de la norma:*

i. $||x|| \geq 0$. ***positividad***

ii. $||x|| = 0$ *sii* $x = 0$. ***no degeneración***

iii. $||\alpha x|| = |\alpha|\ ||x||$ *para todo* $\alpha \in \mathbb{R}$. ***multiplicatividad***

iv. $||x + y|| \leq ||x|| + ||y||$. ***desigualdad triangular***

III. *Propiedades de la distancia:*

i. $d(x, y) \geq 0$. ***positividad***

ii. $d(x, y) = 0$ *sii* $x = y$. ***no degeneración***

iii. $d(x, y) = d(y, x)$. ***simetría***

iv. $d(x, y) \leq d(x, z) + d(z, y)$. ***desigualdad triangular***

IV. *Desigualdad de Cauchy-Schwarz:*

$|\langle x, y\rangle| \leq ||x||\ ||y||$.

Las propiedades **i–iv** del producto interno son casi obvias desde un punto de vista algebraico. Una vez que tenemos la desigualdad de Cauchy-Schwarz, las propiedades de la norma y la distancia se deducen fácilmente.

He aquí una prueba de la ***desigualdad de Cauchy-Schwarz*** (damos otra en la sección de demostraciones de los teoremas). Queremos probar que $|\langle x, y\rangle| \leq \|x\|\ \|y\|$. Los casos $x = 0$ e $y = 0$ son triviales, y por lo tanto podemos suponer que $x \neq 0$ e $y \neq 0$. Si reescalamos x e y expresando $|\langle x, y\rangle| \leq \|x\|\ \|y\|$ en la forma $|\langle x/\|x\|, y/\|y\|\rangle| \leq 1$, podemos suponer que $\|x\| = \|y\| = 1$; así pues, sólo hay que probar que $|\langle x, y\rangle| \leq 1$. De acuerdo con el teorema de Pitágoras, podemos afirmar que

$$||\langle x, y\rangle y||^2 + ||x - \langle x, y\rangle y||^2 = ||x||^2 = 1,$$

porque $\langle x, y\rangle y$ debe ser la proyección del vector x sobre el vector y, como se ve en la figura 1.6-3. Para verificarlo de forma algebraica escribimos

$$\begin{aligned} ||\langle x, y\rangle y||^2 + ||x - \langle x, y\rangle y||^2 &= \langle x, y\rangle^2 ||y||^2 + \langle x - \langle x, y\rangle y, x - \langle x, y\rangle y\rangle \\ &= \langle x, y\rangle^2 ||y||^2 + ||x||^2 - 2\langle x, y\rangle^2 + \langle x, y\rangle^2 ||y||^2 \\ &= ||x||^2 \end{aligned}$$

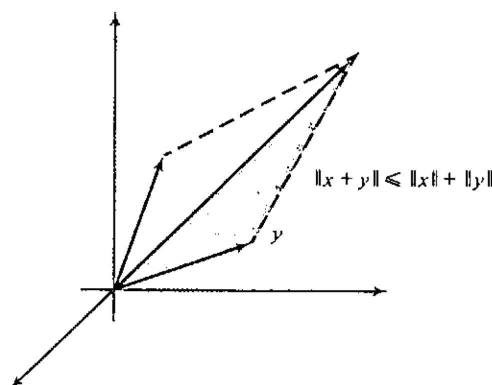


FIGURA 1.6-4 Desigualdad triangular

ya que $\|y\| = 1$. Por consiguiente

$$|\langle x,y\rangle|^2 = \|\langle x,y\rangle y\|^2 \leq \|\langle x,y\rangle y\|^2 + \|x - \langle x,y\rangle y\|^2 = \|x\|^2 = 1,$$

de donde se deduce la desigualdad de Cauchy-Schwarz. Esta demostración también prueba que si en la desigualdad de Cauchy-Schwarz se da la igualdad, entonces x e y son linealmente dependientes (el término $\|x - \langle x,y\rangle y\|^2$ en la última fórmula es cero en este caso, y, por lo tanto, $\|x\| = \langle x,y\rangle y$). Obsérvese que **IV** se deduce sólo a partir de las propiedades **I**.

Muchas de estas propiedades deberían ser evidentes desde un punto de vista geométrico. Por ejemplo, **iv** en **II** y **III** expresa el hecho de que la longitud de un lado de un triángulo es menor o igual que la suma de las longitudes de los otros dos lados (figura 1.6-4).

Un conjunto con una función d que satisface las propiedades del grupo **III** se llama ***espacio métrico***; un espacio vectorial con una norma que obedece las propiedades del grupo **II** se llama ***espacio normado***, y un espacio vectorial con un producto interno que cumple las reglas del grupo **I** es un ***espacio con producto interno***. Veremos estos conceptos en §**1.7** y probaremos que las propiedades **III** se deducen de las **II**, y que las propiedades **II** se deducen de las **I**. Por lo tanto, todo lo que necesitamos probar es que el producto interno propuesto en $\mathbb{R}^n$ satisface las propiedades **I**. (Aclaración: la famosa desigualdad del grupo **IV** se debería llamar ***desigualdad de Cauchy-Bunyakovskii-Schwarz***.)

Generalizando los conceptos de $\mathbb{R}^3$, diremos que x e $y \in \mathbb{R}^n$ son ***ortogonales*** sii $\langle x,y\rangle = 0$. Dos subespacios S y T son ortogonales sii $\langle x,y\rangle = 0$ para todo $x \in S$ e $y \in T$. Si además S y T generan todo $\mathbb{R}^n$, entonces se llaman ***complementos ortogonales***.

Una observación que nos será útil al estudiar sucesiones en $\mathbb{R}^n$ es la siguiente:

1.6.7 Proposición *Si $v = (v_1, \ldots, v_n)$ y $w = (w_1, \ldots, w_n)$ son vectores de $\mathbb{R}^n$ y definimos $\rho(v, w) = \text{máx}\{|v_1 - w_1|, |v_2 - w_2|, \ldots, |v_n - w_n|\}$, entonces*

$$\rho(v, w) \leq \|v - w\| \leq \sqrt{n}\rho(v, w).$$

1.6.8 Ejemplo *Calcúlese la longitud del segmento de recta que une* (1, 1, 1) *y* (3, 2, 0).

Solución Se trata de la longitud del vector (3, 2, 0) – (1, 1, 1) = (2, 1, –1), el cual representa el vector que va de (1, 1, 1) a (3, 2, 0). La longitud es

$$\|(2, 1, -1)\| = \sqrt{2^2 + 1^2 + (-1)^2} = \sqrt{6}. \quad ♦$$

1.6.9 Ejemplo *En $\mathbb{R}^3$, encuéntrese el complemento ortogonal de la recta $x = y = z/2$ (o $x_1 = x_2 = x_3/2$ con otra notación).*

Solución Esta recta, que llamaremos l, es un subespacio unidimensional generado por el vector (1, 1, 2) (véase la figura 1.6-5). El complemento ortogonal es un plano (que pasa por el origen, puesto que es un subespacio) y, por tanto, tiene una ecuación de la forma

$$Ax + By + Cz = 0, \quad \text{es decir,} \quad \langle (A, B, C), (x, y, z) \rangle = 0.$$

Así pues, (A, B, C) es normal al plano; pero (1, 1, 2) es un vector perpendicular al plano, así que el complemento ortogonal que buscamos es el plano $x + y + 2z = 0$. ♦

Ejercicios de §1.6

1. Si $\|x + y\| = \|x\| + \|y\|$, demuéstrese que x e y están en la misma recta que pasa por el origen.

2. ¿Cuál es el ángulo entre (3, 2, 2) y (0, 1, 0)?

3. Hállese el complemento ortogonal del plano generado por (3, 2, 2) y (0, 1, 0) en $\mathbb{R}^3$.

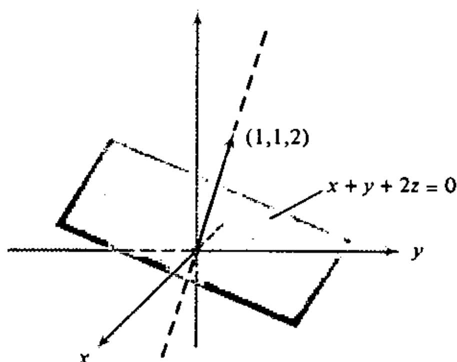


FIGURA 1.6-5 La recta l y su complemento ortogonal en el ejemplo **1.6.10**

4. Descríbanse los conjuntos $B = \{x \in \mathbb{R}^3 \mid \|x\| \leq 3\}$ y $Q = \{x \in \mathbb{R}^3 \mid \|x\| < 3\}$.

5. Hállese la ecuación de la recta que pasa por (1, 1, 1) y (2, 3, 4). ¿Es esta recta un subespacio lineal?

§1.7 Normas, productos internos y métricas

Entre las propiedades de $\mathbb{R}^n$ está su propiedad métrica, que proporciona un modo de medir distancias entre puntos. Un ***espacio métrico*** es un conjunto M con una función $d : M \times M \to \mathbb{R}$ que nos da una manera razonable de medir la distancia entre dos elementos de M. La notación y el lenguaje están diseñados para que recuerden la distancia entre puntos en el espacio euclídeo uni, bi o tridimensional, que nos es tan familiar, pero no para restringirnos a ella.

1.7.1 Definición *Un **espacio métrico** (M, d) es un conjunto M y una función $d : M \times M \to \mathbb{R}$ que cumplen las siguientes propiedades:*

i. $d(x, y) \geq 0$ *para todo* $x, y \in M$. ***positividad***

ii. $d(x, y) = 0$ *si y sólo si* $x = y$. ***no degeneración***

iii. $d(x, y) = d(y, x)$ *para todo* $x, y \in M$. ***simetría***

iv. $d(x, y) \leq d(x, z) + d(z, y)$ *para todo* $x, y, z \in M$. ***desigualdad triangular***

1.7.2 Ejemplos

a. *Ya vimos en* **1.1.6** *que la recta real,* $\mathbb{R}$, *es un espacio métrico con la métrica definida por* $d(x, y) = |x - y|$. *El lector que conozca los números complejos se dará cuenta de que* $\mathbb{C}$ *también es un espacio métrico con* $d(z, w) = |z - w|$; *véase* **§1.8** *para más detalles. De manera similar,* **1.6.6** *muestra que* $\mathbb{R}^n$ *es un espacio métrico (con la métrica habitual).*

b. ***Métrica discreta*** *Sea M cualquier conjunto y sea* $d(x, y) = 0$ *si* $x = y$ *y* $d(x, y) = 1$ *si* $x \neq y$. *Entonces d es una métrica sobre M.*

c. ***Métrica lexicográfica*** *Este ejemplo es útil en combinatoria y en informática. Sea M el conjunto de todas las "palabras", cada una de las cuales consiste en un octeto ordenado de ceros y unos:*

$$w = (w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8) \quad \textit{donde cada "bit"} \quad w_k = 0 \quad o \quad 1.$$

Sea $d(v, w)$ = *el número de lugares en los que v y w son diferentes. Por ejemplo, si* $v = 01100011$ *y* $w = 00110101$, *entonces* $d(v, w) = 4$, *porque v y w difieren en los bits segundo, cuarto, sexto y séptimo. Entonces d es una métrica en M. (Para probar sus propiedades puede ser útil verificar que* $d(v, w) = \sum_{k=1}^{8} |v_k - w_k|$.)

d. ***Métrica acotada*** *Si d es una métrica sobre un conjunto M y* $\rho(x,y)$ *se define mediante* $\rho(x, y) = d(x, y)/(1 + d(x, y))$, *entonces* ρ *también es una métrica sobre M, tal como puede comprobarse (véase el ejercicio 10 del capítulo 1). Esta nueva métrica tiene la propiedad de que* $\rho(x, y) < 1$ *para todo par* $x, y \in M$; *es decir, está acotada por* 1. ♦

Aunque en ocasiones los espacios métricos son espacios vectoriales, esto no siempre ocurre, como lo prueba el espacio discreto. Vamos a centrarnos ahora en conceptos que sólo tienen sentido en espacios vectoriales.

1.7.3 Definición *Un* ***espacio normado*** *(o* ***espacio vectorial normado****)* $(\mathcal{V}, \|\cdot\|)$ *es un espacio vectorial* $\mathcal{V}$ *y una función* $\|\cdot\|: \mathcal{V} \to \mathbb{R}$, *llamada* ***norma****, tal que*

i. $\|v\| \geq 0$ *para todo* $v \in \mathcal{V}$. ***positividad***

ii. $\|v\| = 0$ *si y sólo si* $v = 0$. ***no degeneración***

iii. $\|\lambda v\| = |\lambda| \, \|v\|$ *para todo* $v \in \mathcal{V}$ *y todo escalar* λ. ***multiplicidad***

iv. $\|v + w\| \leq \|v\| + \|w\|$ ***desigualdad triangular***

1.7.4 Ejemplos

a. ***Números reales*** Ya vimos en **1.1.5** que $\mathbb{R}$ es un espacio normado con $\|x\| = |x|$. El lector que conozca los números complejos se percatará de que $\mathbb{C}$ es un espacio normado con $\|z\| = |z|$.

b. ***Norma del taxi*** Considérese el espacio $\mathbb{R}^2$, pero en lugar de la norma usual, definamos $\|(x, y)\|_1 = |x| + |y|$; $\|\cdot\|_1$ es una norma en $\mathbb{R}^2$. El nombre proviene de la distancia asociada. Si $P = (x, y)$ y $Q = (a, b)$, entonces

$$d_1(P, Q) = \|P - Q\|_1 = |x - a| + |y - b|.$$

o sea, la suma de las separaciones vertical y horizontal. Ésta es la distancia que se debe recorrer para llegar de P a Q si siempre se viaja paralelamente a los ejes (que es lo que hace un taxi al recorrer las calles). Véase la figura 1.7-1.

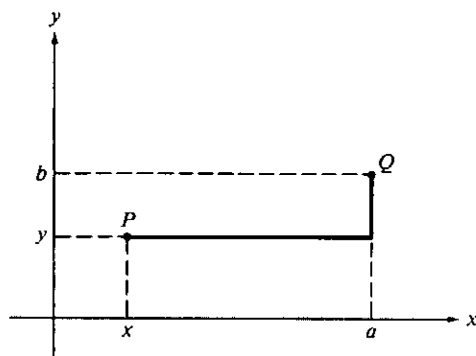


FIGURA 1.7-1 La métrica del taxi

c. ***Norma del supremo*** Sea M el conjunto de todas las funciones reales definidas en el intervalo $[0, 1]$ que están acotadas; es decir,

$$M = \{f : [0, 1] \to \mathbb{R} \mid \text{existe un número } B \text{ que cumple } |f(x)| \le B \text{ para todo } x \in [0, 1]\}.$$

Para cada f en M, $f([0,1])$ es un subconjunto acotado de $\mathbb{R}$ y por lo tanto $\{|f(x)| \mid x \in [0, 1]\}$ también es acotado; en consecuencia, tiene un supremo finito y

$$\|f\|_\infty = \sup\{|f(x)| \mid x \in [0, 1]\}$$

define, por lo tanto, una función $\|\cdot\|_\infty : M \to \mathbb{R}$. El conjunto M es un espacio vectorial y $\|\cdot\|_\infty$ es una norma sobre él. Estudiaremos más profundamente esta norma en el capítulo 5. ♦

Como hicimos notar anteriormente, al hablar sobre $\mathbb{R}^n$, las normas siempre engendran métricas.

1.7.5 Proposición *Si $(V, \|\cdot\|)$ es un espacio vectorial normado y $d(v, w)$ está definida por*

$$d(v, w) = \|v - w\|,$$

entonces d es una métrica sobre V.

En algunas ocasiones, las métricas engendradas por normas proporcionan una medida de la diferencia entre dos vectores. Ya vimos un ejemplo en la métrica del taxi. En el ejemplo del espacio de las funciones acotadas, la métrica es

$$d(f, g) = \|f - g\|_\infty = \sup\{|f(x) - g(x)| \mid 0 \leq x \leq 1\}.$$

Por tanto, la métrica dada por la norma del supremo es la mayor separación vertical entre las gráficas, como en la figura 1.7-2.

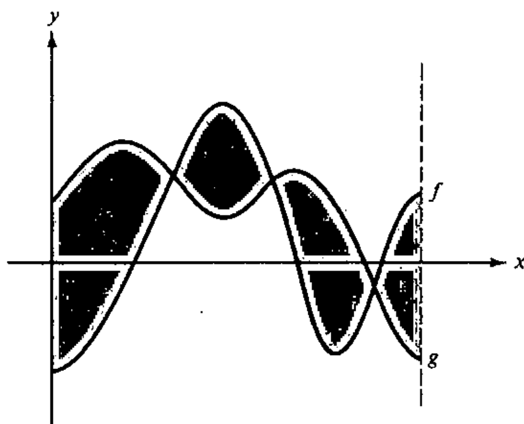


FIGURA 1.7-2 La distancia del supremo entre dos funciones es la máxima distancia vertical entre sus gráficas

Aunque muchas métricas interesantes provienen de normas, no todas lo hacen. Salvo que el espacio vectorial consista sólo en el vector cero, elíjase un vector v diferente de cero y obsérvese que $\|\lambda v\| = |\lambda| \|v\| \to \infty$ cuando $|\lambda| \to \infty$, por lo que $d(\lambda v, 0)$ puede ser arbitrariamente grande. En consecuencia, una norma no puede producir las métricas de los ejemplos **1.7.2b** y **d**.

1.7.6 Definición *Un espacio vectorial real V con una función $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ se llama* ***espacio con producto interno*** *si*

i. $\langle v, v\rangle \geq 0$ *para todo* $v \in V$. ***positividad***

ii. $\langle v, v\rangle = 0$ *si y sólo si* $v = 0$. ***no degeneración***

iii. $\langle \lambda v, w\rangle = \lambda \langle v, w\rangle$ *para todo* $v, w \in V$ *y para todo* $\lambda \in \mathbb{R}$. ***multiplicatividad***

iv. $\langle v, w + u\rangle = \langle v, w\rangle + \langle v, u\rangle$ *para todo* $v, w, u \in V$. ***distributividad***

v. $\langle v, w\rangle = \langle w, v\rangle$ *para todo* $v, w \in V$. ***simetría***

El ejemplo más importante es $\mathbb{R}^n$ con el producto interno usual, pero hay otros.

1.7.7 Ejemplo ***Funciones continuas*** *Sea V el espacio de todas las funciones reales continuas sobre el intervalo* $[0, 1]$. *(La continuidad se estudiará en detalle en el capítulo 4. Por ahora usaremos la noción de manera informal, suponiendo que el lector la conoce de algún curso de cálculo.) Como el lector probablemente sabrá, la suma de funciones continuas es continua (pronto verificaremos esto de manera oficial); por lo tanto, V es un espacio vectorial, que habitualmente se denota* $C([0, 1])$, *y que tiene un gran interés en análisis. Confirmaremos en el capítulo 4 que cualquier función continua en* $[0, 1]$ *es acotada, de donde la norma del supremo definida en el ejemplo* **1.7.4c** *tiene sentido en* $C([0, 1])$. *Veremos en el ejercicio 30 al final de este capítulo que esta norma no se puede asociar con un producto interno. Existe, sin embargo, un producto interno muy interesante en* $C([0, 1])$ *no asociado con la norma del supremo. Dado que el producto de funciones continuas es una función continua y las funciones continuas son integrables (verificaremos estas afirmaciones más adelante, pero por ahora las tomaremos como ciertas), podemos definir un producto interno en* $C([0, 1])$ *mediante*

$$\langle f, g\rangle = \int_0^1 f(x)g(x)dx.$$

Es inmediato (dadas las propiedades básicas de las integrales) verificar que éste es un producto interno. Será muy importante en el capítulo 10 para estudiar las series de Fourier y el análisis de Fourier. ♦

El producto interno, cuando se dispone de él, es una herramienta muy poderosa. Como en $\mathbb{R}^n$, de él proviene la noción de ortogonalidad: dos vectores v y w son

ortogonales si $\langle v, w\rangle = 0$. También nos proporciona una de las desigualdades más útiles en análisis: la desigualdad de Cauchy-Schwarz (o de Cauchy-Bunyakovskii-Schwarz). Antes de volver sobre ella, recopilaremos algunos resultados sencillos pero útiles.

1.7.8 Proposición *Si $\langle\cdot,\cdot\rangle$ es un producto interno sobre un espacio vectorial real V, entonces*

i. $\langle\lambda v + \mu w, u\rangle = \lambda\langle v, u\rangle + \mu\langle w, u\rangle$

ii. $\langle u, \lambda v + \mu w\rangle = \lambda\langle u, v\rangle + \mu\langle u, w\rangle$

iii. $\langle v, \lambda w\rangle = \lambda\langle v, w\rangle$

iv. $\langle 0, w\rangle = \langle w, 0\rangle = 0$.

La demostración es inmediata y se deja al lector.

1.7.9 Desigualdad de Cauchy-Schwarz *Si $(V, \langle\cdot,\cdot\rangle)$ es un espacio con producto interno, entonces*

$$|\langle v, w\rangle| \leq \sqrt{\langle v, v\rangle}\ \sqrt{\langle w, w\rangle}$$

para todo v y w en V.

¡La demostración que dimos en la sección anterior también es válida en este contexto abstracto! Esta desigualdad, además de tener otros usos, es la clave para probar que todo producto interno engendra una norma.

1.7.10 Proposición *Si $(V, \langle\cdot,\cdot\rangle)$ es un espacio con producto interno y $\|\cdot\|$ está definida para cada v en V por*

$$\|v\| = \sqrt{\langle v, v\rangle},$$

entonces $\|\cdot\|$ es una norma sobre V.

El punto clave es que la desigualdad triangular se puede obtener como una consecuencia sencilla de la desigualdad de Cauchy-Schwarz:

$$\begin{aligned}\|v + w\|^2 &= \langle v + w, v + w\rangle = \|v\|^2 + 2\langle v, w\rangle + \|w\|^2 \\ &\leq \|v\|^2 + 2|\langle v, w\rangle| + \|w\|^2 \\ &\leq \|v\|^2 + 2\|v\|\ \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2,\end{aligned}$$

y, por lo tanto, $\|v + w\| \leq \|v\| + \|w\|$.

Ejercicios de §1.7

1. Encuéntrese $d(f, g)$ en $C([0, 1])$, donde $f(x) = 1$ y $g(x) = x$, usando tanto la norma del supremo como la norma dada en el ejemplo **1.7.7.**

2. Identifíquese el espacio de los polinomios de grado $n - 1$ de una variable real con $\mathbb{R}^n$. Calcúlese $d(f,g)$, donde $f(x) = 1$ y $g(x) = x$, con la métrica euclídea correspondiente.

3. Dótese del producto interno $\langle f, g\rangle = \int_0^1 f(x)g(x)dx$ a $C([0, 1])$. Verifíquese la desigualdad de Cauchy-Schwarz para $f(x) = 1$ y $g(x) = x$.

4. Usando el producto interno del ejercicio 3, verifíquese la desigualdad triangular para $f(x) = x$ y $g(x) = x^2$.

5. Demuéstrese que $\|\cdot\|_\infty$ no es la norma definida por la métrica del ejemplo **1.7.7.**

§1.8 Los números complejos

Nota: El material que sigue es importante para el análisis en general, pero no es esencial para este libro, excepto para el capítulo 10. Estos temas pueden estudiarse ampliamente en cursos de variable compleja; véase, por ejemplo, *Basic Complex Analysis* de J. Marsden y M. Hoffman, 2ª ed., W.H. Freeman, Nueva York, 1987.

Los números reales $\mathbb{R}$ se crearon con objeto de tener una recta continua de números, adecuada para las necesidades de la geometría y el cálculo. Sin embargo, aún quedan algunos problemas: mientras que hemos proporcionado una raíz cuadrada de 2, con objeto de que la ecuación $x^2 - 2 = 0$ tenga solución, la ecuación $x^2 + 2 = 0$ no tiene ninguna solución. Los *números complejos* aportan las soluciones de este tipo de ecuaciones. El conjunto de los números complejos, $\mathbb{C}$, se creó siguiendo la clave que proporciona la fórmula cuadrática, la cual establece que la ecuación $ax^2 + bx + c = 0$, con $a \neq 0$, tiene las soluciones $x = -b \pm \sqrt{b^2 - 4ac}/2a$. Para la ecuación $x^2 + x - 6 = 0$ se obtienen las soluciones $x = 2$ y $x = -3$. En el caso de la ecuación $x^2 + 2x + 5 = 0$, la fórmula sugiere que $x = -1 \pm 2\sqrt{-1}$, por lo que a veces en el colegio se nos enseña a concluir que no hay solución. Pero supongamos que tratamos de hallar una simplemente como un juego matemático. Si estos valores de x se insertan en la ecuación y se manipulan según las reglas de la aritmética, con la regla adicional de que $(\sqrt{-1})^2 = -1$, comprobaremos que, en efecto, "resuelven" la ecuación. Como la fórmula cuadrática da valores reales siempre que $b^2 - 4ac \geq 0$ y sugiere las "soluciones" $z = (-b/2a) \pm$

$(\sqrt{4ac-b^2}/2a)\sqrt{-1}$ cuando $b^2 - 4ac < 0$, ello nos lleva a "jugar" con objetos $z = x + y\sqrt{-1}$, donde x e y son reales. Usaremos el símbolo i para representar $\sqrt{-1}$, con objeto de simplificar la notación. Así pues, manejaremos las expresiones $x + iy$ según las reglas de la aritmética, con la excepción de que $i^2 = -1$. Si $x, y, a, b \in \mathbb{R}$, entonces hallamos que

$$\begin{aligned}(x + iy) + (a + ib) &= (x + a) + i(y + b)\\ (x + iy) + (0 + i0) &= x + iy\\ (x + iy) + (-x - iy) &= 0 + i0\\ (x + iy) \cdot (a + ib) &= (ax - by) + i(bx + ay)\\ (x + iy) \cdot (1 + i0) &= x + iy.\end{aligned}$$

¿Podemos dividir? Al menos de una manera formal, podemos escribir

$$\frac{1}{x+iy} = \frac{1}{x+iy}\frac{x-iy}{x-iy} = \frac{x}{x^2+y^2} + i\frac{-y}{x^2+y^2}.$$

Si por lo menos uno de los valores x o y es distinto de 0, entonces ésta es una expresión válida en nuestro sistema. La multiplicación directa confirma que

$$\left(\frac{x}{x^2+y^2} + i\frac{-y}{x^2+y^2}\right) \cdot (x + iy) = (1 + i0).$$

Daremos una base lógica y una realización geométrica de los números complejos identificando la construcción $x + iy$ con el par ordenado (x, y) y por tanto con un punto en el plano euclídeo de la geometría analítica. El eje horizontal, que consta de los puntos $(x, 0)$, se identifica, pues, con el conjunto de los números reales mediante x " = " $x + i0$ " = " $(x, 0)$. Se le conoce como el ***eje real***. El eje vertical, que consta de los puntos $(0, y)$, se conoce tradicionalmente como el ***eje imaginario***. El punto $(0, 1)$ se identifica con i y se le suele llamar la ***unidad imaginaria***. Las reglas de suma y multiplicación se convierten en definiciones:

$$\begin{aligned}(x, y) + (a, b) &= (x + a, y + b)\\ (x, y) \cdot (a, b) &= (ax - by, bx + ay).\end{aligned}$$

Verificar que el plano con estas operaciones, con $(0, 0)$ como el neutro para la suma y $(1, 0)$ como identidad multiplicativa, es un cuerpo es una labor tediosa, pero rutinaria (satisface las propiedades de la 1 a la 10 de los axiomas de cuerpo dados en §**1.1**). El recíproco de un punto diferente de cero está dado, como sugiere nuestra fórmula para $1/(x + iy)$, por

$$(x, y)^{-1} = \left(\frac{x}{x^2+y^2}, -\frac{y}{x^2+y^2}\right).$$

Más aún, los números reales se pueden identificar con el eje horizontal mediante $x \leftrightarrow (x, 0)$, y todo concuerda: $x + y \leftrightarrow (x, 0) + (y, 0)$ y $xy \leftrightarrow (xy, 0)$. Por último, verificamos que $(0, 1) \cdot (0, 1) = (-1, 0) \leftrightarrow -1$, y por lo tanto tenemos la raíz cuadrada de -1.

El cuerpo resultante se denomina cuerpo de los ***números complejos*** y se denota $\mathbb{C}$. Sin embargo, $\mathbb{C}$ no puede ser un cuerpo ordenado. En un cuerpo ordenado sabemos que $z^2 \geq 0$ para todo elemento z y que $-1 < 0$. Pero en $\mathbb{C}$ tenemos un elemento, i, tal que $i^2 = -1$. En consecuencia, *no hay manera de introducir un orden en* $\mathbb{C}$ *que tenga todas las propiedades que definen un cuerpo ordenado.*

La identificación de $\mathbb{C}$ con el plano $\mathbb{R}^2$ nos da una interpretación geométrica. La suma es la misma que la suma de vectores en el plano y cumple la ley del paralelogramo, como en la figura 1.8-1.

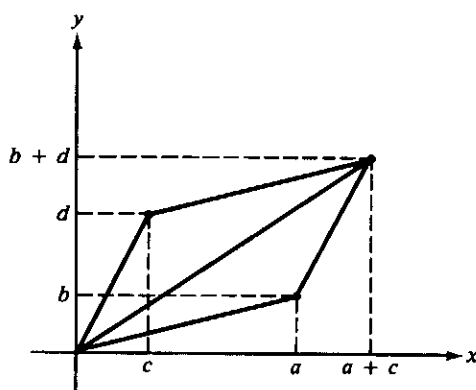


FIGURA 1.8-1 Los números complejos se suman como vectores, lo que ilustra $(a + bi) + (c + di) = (a + c) + (b + d)i$

El significado geométrico de la multiplicación queda claro si usamos coordenadas polares. Supongamos que

$$z = (r\cos\theta, r\operatorname{sen}\theta) = r\cos\theta + ir\operatorname{sen}\theta$$
$$w = (\rho\cos\varphi, \rho\operatorname{sen}\varphi) = \rho\cos\varphi + i\rho\operatorname{sen}\varphi.$$

Entonces

$$\begin{aligned} z \cdot w &= (r\cos\theta + ir\operatorname{sen}\theta)(\rho\cos\varphi + i\rho\operatorname{sen}\varphi) \\ &= r\rho[(\cos\theta\cos\varphi - \operatorname{sen}\theta\operatorname{sen}\varphi) + i(\operatorname{sen}\theta\cos\varphi + \cos\theta\operatorname{sen}\varphi)] \\ &= r\rho[\cos(\theta + \varphi) + i\operatorname{sen}(\theta + \varphi)] = [r\rho\cos(\theta + \varphi), r\rho\operatorname{sen}(\theta + \varphi)]. \end{aligned}$$

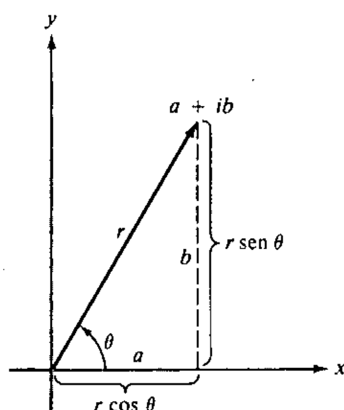


FIGURA 1.8-2 Representación de los números complejos mediante coordenadas polares

Para multiplicar z por w multiplicamos las magnitudes de los vectores y sumamos los ángulos polares. Si pensamos que w está fijo, multiplicarlo por z corresponde a girarlo un ángulo θ y alargarlo un factor r.

r es la ***magnitud*** o ***valor absoluto*** de z y se denota z.

θ es el ***argumento*** de z y se denota arg(z).

Obsérvese que si $z = x + iy$, con x e y reales, entonces $|z| = \sqrt{x^2 + y^2}$, que es la distancia en el plano entre el punto z y el origen. Los cálculos con estas cantidades se simplifican si introducimos el ***complejo conjugado*** de z. Éste se denota $\bar{z}$ y se define como $\bar{z} = x - iy$. La coordenada horizontal x se denomina ***parte real*** de z y se denota Re z o Re(z), mientras que y se denomina ***parte imaginaria*** y se denota Im z o Im(z).

Es útil tener una definición de e^z donde z pueda ser un número complejo. Como queremos que e^a coincida con la definición usual para a real y queremos que $e^{a+bi} = e^a \cdot e^{bi}$, sólo necesitamos definir $e^{i\theta}$ para θ real. *Definimos* $e^{i\theta} = \cos\theta + i \operatorname{sen}\theta$, con lo que $e^{a+bi} = e^a(\cos b + i \operatorname{sen} b)$. Entonces, como $\cos 0 + i \operatorname{sen} 0 = 1$, tenemos $e^{a+0i} = e^a$ y, por lo tanto, nuestra definición coincide con la definición usual en el caso en que a sea real. El lector puede también verificar (véase el ejercicio 7a de esta sección) que $e^{z_1+z_2} = e^{z_1} \cdot e^{z_2}$.

Los números complejos se representan (como ya se ha dicho) mediante puntos en $\mathbb{R}^2$. Usando coordenadas polares, podemos, pues, escribir

$$z = re^{i\theta} = r\cos\theta + ir\operatorname{sen}\theta,$$

donde $r = |z|$ y $\theta = \arg z$, el argumento de z; véase la figura 1.8-2.

1.8.1 Proposición *Supongamos que $z = x + iy$, $w = u + iv \in \mathbb{C}$, y $x, y, u, v \in \mathbb{R}$. Entonces*

i. $|z| \geq 0$.

ii. $|z| = 0 \Leftrightarrow z = 0$.

iii. $|zw| = |z||w|$.

iv. $|z + w| \leq |z| + |w|$.

v. $\overline{zw} = \bar{z}\,\bar{w}$.

vi. $\overline{z + w} = \bar{z} + \bar{w}$.

vii. $\overline{(1/z)} = 1/\bar{z}$.

viii. $|z|^2 = z\bar{z}$.

ix. $|z|^2 = x^2 + y^2$.

x. $\text{máx}(|x|, |y|) \leq |z| \leq \sqrt{2}\,\text{máx}(|x|, |y|)$.

1.8.2 Ejemplos

a. *Pruébese que $1/i = -i$ y que $1/(i + 1) = (i - 1)/2$.*

Solución En primer lugar,

$$\frac{1}{i} = \frac{1}{i} \cdot \frac{-i}{-i} = -i,$$

porque $i \cdot -i = -(i^2) = -(-1) = 1$. También,

$$\frac{1}{i+1} = \frac{1}{i+1}\frac{1-i}{1-i} = \frac{1-i}{2},$$

ya que $(1 + i)(1 - i) = 1 + 1 = 2$.

b. *Hállense las partes real e imaginaria de $(z + 2)/(z - 1)$, donde $z = x + iy$.*

Solución

$$\begin{aligned}\frac{z+2}{z-1} &= \frac{(x+2)+iy}{(x-1)+iy} = \frac{(x+2)+iy}{(x-1)+iy} \cdot \frac{(x-1)-iy}{(x-1)-iy} \\ &= \frac{(x+2)(x-1)+y^2+i[y(x-1)-y(x+2)]}{(x-1)^2+y^2}.\end{aligned}$$

Por lo tanto,

$$\operatorname{Re}\left[\frac{z+2}{z-1}\right]=\frac{x^2+x-2+y^2}{(x-1)^2+y^2} \quad \text{e} \quad \operatorname{Im}\left[\frac{z+2}{z-1}\right]=\frac{-3y}{(x-1)^2+y^2}.$$

c. *Demuéstrese que* $\overline{\left[\frac{(3+7i)^2}{8+6i}\right]}=\frac{(3-7i)^2}{8-6i}.$

Solución La clave en este caso es que no es necesario desarrollar $(3+7i)^2/(8+6i)$ si aplicamos las propiedades desarrolladas en el texto, a saber $\overline{z^2}=\bar{z}^2$, y $\overline{z/z'}=\bar{z}/\bar{z}'$. Así.

$$\overline{\left[\frac{(3+7i)^2}{8+6i}\right]}=\frac{\overline{(3+7i)^2}}{\overline{8+6i}}=\frac{\overline{(3+7i)}^2}{8-6i}=\frac{(3-7i)^2}{8-6i}.$$

d. *Si* $|z|=1$ *pruébese que* $\left|\frac{az+b}{\bar{b}z+\bar{a}}\right|=1$ *para todo par de complejos a y b.*

Solución Como $|z|=1$, tenemos $z=(\bar{z})^{-1}$. Por lo tanto,

$$\frac{az+b}{\bar{b}z+\bar{a}}=\frac{az+b}{\bar{b}+\bar{a}\,\bar{z}}\cdot\frac{1}{z}.$$

Por las propiedades del valor absoluto y porque $|z|=1$,

$$\left|\frac{az+b}{\bar{b}z+\bar{a}}\right|=\left|\frac{az+b}{\overline{az+b}}\right|\cdot\frac{1}{|z|}=1$$

ya que $|az+b|=|\overline{az+b}|=|\bar{a}\,\bar{z}+\bar{b}|$.

e. *Pruébese que el máximo valor absoluto de* z^2+1 *sobre el disco unidad* $|z|\le 1$ *es* 2.

Solución Por la desigualdad triangular, $|z^2+1|\le|z^2|+1=|z|^2+1\le 1^2+1=2$ ya que $|z|=\le 1$, luego $|z^2+1|$ no es mayor que 2 en el disco. Como el valor 2 se alcanza en $z=1$, el máximo es 2. ♦

Las sucesiones de números complejos no presentan dificultades nuevas. Como el valor absoluto de $\mathbb{C}$ satisface las mismas propiedades formales que en $\mathbb{R}$ y es, de hecho, una norma, no existe dificultad para definir la ***convergencia*** de una sucesión de números complejos $z_1, z_2, z_3, \ldots$ a un número z como

$$\lim_{k\to\infty} z_k = z \textit{ sii para todo } \varepsilon > 0 \textit{ existe un número } N$$
$$\textit{tal que } |z_k - z| < \varepsilon \textit{ siempre que } k \geq N.$$

La aritmética de sucesiones es igual que la de $\mathbb{R}$. Por ejemplo, si $z_n \to z$, $w_n \to w$ y λ es constante, $z_n + w_n \to z + w$, $\lambda z_n \to \lambda z$ y $z_n w_n \to zw$. Si $z_n \neq 0$ y $z \neq 0$, entonces $w_n/z_n \to w/z$. Usando las desigualdades

$$|\mathrm{Re}(z) - \mathrm{Re}(w)| \leq |z - w| \leq |\mathrm{Re}(z) - \mathrm{Re}(w)| + |\mathrm{Im}(z) - \mathrm{Im}(w)|$$

y

$$|\mathrm{Im}(z) - \mathrm{Im}(w)| \leq |z - w| \leq |\mathrm{Re}(z) - \mathrm{Re}(w)| + |\mathrm{Im}(z) - \mathrm{Im}(w)|,$$

obtenemos:

1.8.3 Proposición $z_n \to z$ *en* $\mathbb{C}$ *sii* $\mathrm{Re}(z_n) \to \mathrm{Re}(z)$ *e* $\mathrm{Im}(z_n) \to \mathrm{Im}(z)$ *en* $\mathbb{R}$.

Como en el caso de los números reales, una sucesión acotada debe tener una subsucesión convergente, y de esto se deduce que $\mathbb{C}$ es ***completo***. Por lo tanto, *una sucesión en* $\mathbb{C}$ *converge a un punto en* $\mathbb{C}$ *si y sólo si es una* ***sucesión de Cauchy***. Generalizaremos esto a $\mathbb{R}^n$ en §**2.7**.

En algunas ocasiones necesitamos un producto interno en un espacio vectorial sobre $\mathbb{C}$ en lugar de sobre $\mathbb{R}$. Para ello necesitamos algunos cambios; por ejemplo, si $\langle v, v\rangle = \|v\|^2$, entonces $\langle v, v\rangle$ debe ser *real*, así que $\langle v, v\rangle = \overline{\langle v, v\rangle}$. Esto hace plausible que el cambio apropiado sea modificar la condición de simetría v en la definición de producto interno (definición **1.7.6**).

1.8.4 Definición *Un espacio vectorial complejo* $\mathcal{V}$ *con una función* $\langle\cdot,\cdot\rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{C}$ *se denomina* ***espacio complejo con producto interno*** *si*

i. $\langle v, v\rangle \geq 0$ *para todo* $v \in \mathcal{V}$. ***positividad***

ii. $\langle v, v\rangle = 0$ *si y sólo si* $v = 0$. ***no degeneración***

iii. $\langle \lambda v, w\rangle = \lambda\langle v, w\rangle$ *para todo* $v, w \in \mathcal{V}$ *y* $\lambda \in \mathbb{R}$. ***multiplicatividad***

iv. $\langle v, w+u\rangle = \langle v, w\rangle + \langle v, u\rangle$ *para todo* $v, w, u, \in \mathcal{V}$. ***distributividad***

v. $\langle v, w\rangle = \overline{\langle w, v\rangle}$ *para todo* $v, w \in \mathcal{V}$. ***simetría hermítica***

Por ejemplo, si $\mathbb{C}^n$ es el espacio vectorial complejo de las n-uplas ordenadas de números complejos, es decir, si

$$\mathbb{C}^n = \{(z_1, z_2, \ldots, z_n) \mid z_k \in \mathbb{C} \text{ para cada } k\}$$

y $\langle z, w\rangle$ se define para $z = (z_1, \ldots, z_n)$ y $w = (w_1, \ldots, w_n)$ mediante

$$\langle z, w\rangle = \sum_{k=1}^{n} z_k \overline{w}_k,$$

entonces $\langle\cdot, \cdot\rangle$ es un producto interno complejo en $\mathbb{C}^n$. No debe sorprendernos que la mayoría de las propiedades sigan valiendo sin cambios esenciales, excepto alguna conjugación compleja de vez en cuando. La parte **ii** de la proposición **1.7.8** se cambia por **ii**′ $\langle u, \lambda v + \mu w\rangle = \overline{\lambda}\langle u, v\rangle + \overline{\mu}\langle u, w\rangle$, y el caso especial **iii** se cambia por **iii**′ $\langle v, \lambda w\rangle = \lambda\langle v, w\rangle$. En consecuencia, el producto interno (se suele llamar forma lineal ***hermítica*** o ***sesquilineal***) es lineal en la primera variable y lineal conjugado en la segunda. Una nota de advertencia: algunos autores, en particular los de textos de física, usan otra convención, con la forma lineal en la segunda variable y lineal conjugada en la primera: $\langle v, \lambda w\rangle = \lambda\langle v, w\rangle$ y $\lambda\langle v, w\rangle = \overline{\lambda}\langle v, w\rangle$. Esto no provoca ninguna diferencia importante si se es consistente.

El enunciado de la desigualdad de Cauchy-Schwarz sigue siendo válido en espacios vectoriales complejos, pero la demostración es más engañosa: hay que "rotar" primero los vectores mediante la multiplicación por un escalar complejo de norma 1 con objeto de que los coeficientes apropiados usados en la prueba sean reales. Como antes, la ecuación $\|z\|^2 = \langle z, z\rangle$ define una norma sobre $\mathcal{V}$. El razonamiento para la desigualdad triangular es el siguiente:

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\
&= \|x\|^2 + \langle x, y\rangle + \overline{\langle x, y\rangle} + \|y\|^2 = \|x\|^2 + 2\mathrm{Re}(\langle x, y\rangle) + \|y\|^2 \\
&\leq \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \quad \text{(por Cauchy-Schwarz)} \\
&\leq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.
\end{aligned}$$

Ejercicios de §1.8

1. Exprésense los siguientes números complejos en la forma $a + ib$:

a. $(2 + 3i) + (4 + i)$

b. $(2 + 3i)/(4 + i)$

c. $1/i + 3/(1 + i)$

2. Hállense las partes real e imaginaria de los siguientes complejos, donde $z = x + iy$:

a. $(z + 1)/(2z - 5)$

b. z^3

3. ¿Es verdad que $\mathrm{Re}(zw) = (\mathrm{Re}\ z)(\mathrm{Re}\ w)$?

4. ¿Cuál es el complejo conjugado de $(8-2i)^{10}/(4 + 6i)^5$?

5. ¿Se cumple $z^2 = |z|^2$? De ser cierto, pruébese la igualdad. Si no, ¿para qué valores de z es cierta?

6. Suponiendo que $|z| = 1$ o $|w| = 1$ y que $\bar{z}w \neq 1$, pruébese que

$$\left|\frac{z - w}{1 - \bar{z}w}\right| = 1.$$

7. Pruébese que

a. $e^{z_1} \cdot e^{z_2} = e^{z_1 + z_2}$.

b. $e^z \neq 0$ para cualquier número complejo z.

c. $|e^{i\theta}| = 1$ para cualquier número real θ.

d. $(\cos\theta + i\ \mathrm{sen}\ \theta)^n = \cos n\theta + i\ \mathrm{sen}\ n\theta$.

8. Demuéstrese que

a. $e^z = 1$ sii $z = k2\pi i$ para algún entero k.

b. $e^{z_1} = e^{z_2}$ sii $z_1 - z_2 = k2\pi i$ para algún entero k.

9. ¿Para qué valores de z converge la sucesión $z_n = nz^n$?

Demostraciones de los teoremas del capítulo 1

1.1.2 Proposición *En un cuerpo ordenado son válidas las siguientes propiedades:*

ii. ***Unicidad de los inversos*** *Si $a + x = 0$, entonces $x = -a$. Si $ax = 1$, entonces $x = a^{-1}$.*

vi. $0 \cdot x = 0$ *para todo* x.

viii. $-x = (-1)x$ *para todo* x.

xiii. $0 < 1$.

xiv. $x^2 \geq 0$ *para todo* x.

Demostración Para probar **ii**, supongamos que $a + x = 0$. Entonces

$$-a = -a + 0 = -a + (a + x) = (-a + a) + x = 0 + x = x,$$

y así $-a = x$ como se afirmaba. De la misma manera, si $ax = 1$, entonces $a^{-1} = a^{-1} \cdot 1 = a^{-1}(ax) = (a^{-1}a)x = 1x = x$.

Para **vi:** $0 \cdot x = (0 + 0)x = 0 \cdot x + 0 \cdot x$, y así

$$\begin{aligned} 0 &= 0 \cdot x + (-0 \cdot x) = (0 \cdot x + 0 \cdot x) + (-0 \cdot x) \\ &= 0 \cdot x + (0 \cdot x + (-0 \cdot x)) \\ &= 0 \cdot x + 0 = 0 \cdot x. \end{aligned}$$

Para **viii:**

$$\begin{aligned} x + (-1) \cdot x &= 1 \cdot x + (-1) \cdot x \\ &= (1 + (-1)) \cdot x \\ &= 0 \cdot x = 0 \end{aligned}$$

por **vi**. Por lo tanto, $(-1) \cdot x = -x$ por **ii**.

Para **xiii**, supongamos que $1 \leq 0$. Entonces $1 + (-1) \leq 0 + (-1)$, y así $0 \leq -1$. Podemos entonces usar la propiedad 16: como $0 \leq -1$ y $0 \leq -1$, obtenemos $0 \leq (-1) \cdot (-1) = -(-1) = 1$. Por lo tanto, $1 \leq 0$ y $0 \leq 1$, y así $1 = 0$ por la propiedad 12, en contradicción con la propiedad 10.

Para probar **xiv**, consideremos dos casos: si $x \geq 0$, entonces $x^2 = x \cdot x \geq 0$, por el axioma 16; si $x < 0$, entonces $x^2 = (-(-x))(-(-x)) = (-1)^2(-x)^2$, por **vii** y **viii**, pero $(-1)^2 = 1$, porque $0 = (-1)(-1 + 1) = (-1)^2 + (-1) \cdot 1 = (-1)^2 - 1$, luego $x^2 = (-x)^2 \geq 0$. ■

1.1.7 Proposición *$\leq$ es una buena ordenación de* **N**.

Demostración Supongamos que $S \subset \mathbf{N}$ es un conjunto sin elemento mínimo. Sea $T = \mathbf{N} \backslash S$. Usaremos la inducción para demostrar que $T = \mathbf{N}$ y, por tanto, $S = \varnothing$. Si 0 estuviera en S sería el elemento mínimo, luego $0 \in T$. En lugar de abordar T de manera directas, consideremos $T_0 = \{n \in \mathbf{N} \mid \{0, 1, 2, \ldots, n\} \subset T\}$; como $T_0 \subset T$, lograremos nuestro objetivo si probamos que $T_0 = \mathbf{N}$. El argumento anterior prueba que $0 \in T_0$. Supongamos que $k \in T_0$; entonces $\{0, 1, 2, \ldots, k\} \subset T$. Si $k + 1$ estuviera en S sería el elemento mínimo de S, porque todos los números naturales menores que $k + 1$ están en T. De ahí que $k + 1$ debe estar en T. Como $\{0, 1, 2, \ldots, k, k + 1\} \subset T$ tenemos $k + 1 \in T_0$. Por lo tanto, $T_0 = \mathbf{N}$ por inducción. ■

1.2.2 Lema del sandwich *Supongamos que $x_n \to L$, $y_n \to L$ y que $x_n \leq z_n \leq y_n$ para todo n (es suficiente suponer que existe un N_0 tal que $x_n \leq z_n \leq y_n$ siempre que $n > N_0$). Entonces $z_n \to L$.*

Demostración En efecto, sea $\varepsilon > 0$. Entonces:

1. Existe N_0 tal que $x_n \leq z_n \leq y_n$ siempre que $n \geq N_0$,
2. Existe N_1 tal que $|x_n - L| < \varepsilon$ siempre que $n \geq N_1$, y
3. Existe N_2 tal que $|y_n - L| < \varepsilon$ siempre que $n \geq N_2$.

Eligiendo $N = \text{máx}(N_0, N_1, N_2)$, tendremos que siempre que $n \geq N$, $-\varepsilon < x_n - L \leq z_n - L \leq y_n - L < \varepsilon$, y así $|z_n - L| < \varepsilon$, como se pedía. ■

1.2.5 Proposición *Si x_n es una sucesión en $\mathbb{R}$ tal que $x_n \to x$ y $x_n \to y$, entonces $x = y$.*

Demostración x_n converge a x y a y. Escribamos

$$|x - y| = |x - x_n + x_n - y| \leq |x - x_n| + |x_n - y|$$

por la desigualdad triangular. Si $|x - y| > 0$, entonces, usando $|x - y|/2$ como nuestra ε, podemos elegir N de tal manera que tanto $|x - x_n| < |x - y|/2$ como $|x_n - y| < |x - y|/2$ si $n \geq N$, de donde podemos concluir que $|x - y| < |x - y|$, lo cual no puede ocurrir. Por lo tanto, $|x - y| = 0$ y así $x = y$. ■

1.2.6 Proposición *Una sucesión convergente está acotada.*

Demostración Si $x_n \to x$, existe N tal que $|x_n - x| < 1$ siempre que $n \geq N$. Por lo tanto, $|x_n| \leq |x| + 1$ cuando $n \geq N$. Si definimos $M = \text{máx}\{|x|, |x_1|, |x_2|, \ldots, |x_{n-1}|\} + 1$, obtenemos $|x_n| \leq M$ para todo n. ■

1.2.7 Teorema del límite para sucesiones *Supongamos que $x_n \to x$, $y_n \to y$ y λ es una constante. Entonces*

i. $x_n + y_n \to x + y$.

ii. $\lambda x_n \to \lambda x$.

iii. $x_n y_n \to xy$.

iv. *Si $y_n \neq 0$ e $y \neq 0$, entonces $x_n/y_n \to x/y$.*

Demostración Los argumentos muestran cómo se usa la desigualdad triangular.

i. Sea $\varepsilon > 0$. Elijamos N_1 y N_2 tales que $|x_n - x| < \varepsilon/2$ cuando $n \geq N_1$ y $|y_n - y| < \varepsilon/2$ cuando $n \geq N_2$. Si $N = \text{máx}(N_1, N_2)$, entonces para todo $n \geq N$, tendremos

$$\begin{aligned} |(x_n + y_n) - (x + y)| = |(x_n - x) + (y_n - y)| &\leq |x_n - x| + |y_n - y| \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon, \end{aligned}$$

y por lo tanto $x_n + y_n \to x + y$.

ii. Éste es un caso especial de **iii** en el que y_n es una sucesión constante; en consecuencia, es suficiente demostrar **iii.**

iii. Primero escribimos

$$\begin{aligned} |x_n y_n - xy| &= |x_n y_n - x_n y + x_n y - xy| \\ &\leq |x_n y_n - x_n y| + |x_n y - xy| \\ &= |x_n|\,|y_n - y| + |y|\,|x_n - x|. \end{aligned}$$

Como la sucesión x_n converge, está acotada, esto es, existe una constante M tal que $|x_n| < M$ para todo n. Elijamos N_1 y N_2 de tal modo que $|x_n - x| < \varepsilon/[2(|y| + 1)]$

cuando $n \geq N_1$ y que $|y_n - y| < \varepsilon/[2(M+1)]$ cuando $n \geq N_2$. Sea $N = \text{máx}(N_1, N_2)$. Si $n \geq N$, entonces

$$
\begin{aligned}
|x_n y_n - xy| &\leq |x_n|\,|y_n - y| + |y|\,|x_n - x| \\
&\leq \frac{M\varepsilon}{2(M+1)} + \frac{|y|\varepsilon}{2(|y|+1)} \\
&\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,
\end{aligned}
$$

y por lo tanto $x_n y_n \to xy$.

iv. Si $y \neq 0$, entonces $|y| > 0$, con lo cual existe N_1 tal que $|y_n - y| < |y|/2$ cuando $n \geq N_1$, así que $|y_n| > |y|/2$. Para n así elegida tendremos

$$
\begin{aligned}
\left|\frac{x_n}{y_n} - \frac{x}{y}\right| &= \left|\frac{x_n y - x y_n}{y_n y}\right| = \left|\frac{(x_n - x)y + x(y - y_n)}{y_n y}\right| \\
&\leq \frac{|y|\,|x_n - x|}{|y_n y|} + \frac{|x|\,|y - y_n|}{|y_n y|} \\
&\leq \frac{2}{|y|}\,|x_n - x| + \frac{2|x|}{|y|^2}\,|y_n - y|.
\end{aligned}
$$

Ahora elijamos N_2 y N_3 de tal modo que $|x_n - x| < \varepsilon\,|y|/4$ siempre que $n \geq N_2$, y que $|y_n - y| < \varepsilon\,|y|^2/4|x|$ siempre que $n \geq N_3$. Si $N = \text{máx}(N_1, N_2, N_3)$ y $n \geq N$, entonces

$$
\left|\frac{x_n}{y_n} - \frac{x}{y}\right| < \frac{2}{|y|}\,\frac{\varepsilon|y|}{4} + \frac{4|x|}{|y|^2}\,\frac{\varepsilon|y|^2}{4|x|} = \varepsilon.
$$

y por lo tanto $x_n/y_n \to x/y$. ■

1.2.11 Proposición *Los cuerpos ordenados completos son arquimedianos.*

Demostración Sea F un cuerpo ordenado completo y consideremos $x \in$ F. Debemos probar que existe un entero n tal que $x < n$. Si no existiera tal n, sería una cota superior de la sucesión monótona 1, 2, 3, . . . , que, por lo tanto, convergería a algún valor, digamos y, por la propiedad de la sucesión monótona. Afirmamos que esto no puede ocurrir, ya que, de lo contrario, para cualquier $\varepsilon > 0$, existiría un N tal que si $n \geq N$,

$$
1 = |n + 1 - n| \leq |n + 1 - y| + |y - n| < 2\varepsilon,
$$

por la desigualdad triangular. Esto conduce a una contradicción si $\varepsilon < 1/2$. ■

1.2.17 Proposición *$\mathbb{Q}$ es denso en $\mathbb{R}$. Es decir,*

i. *Si x e y están en $\mathbb{R}$ y $x < y$, existe $r \in \mathbb{Q}$ tal que $x < r < y$.*

ii. *Si $x \in \mathbb{R}$ y $\varepsilon > 0$, entonces existe $r \in \mathbb{Q}$ tal que $x - r < \varepsilon$.*

Demostración La demostración ilustra la relación que hay entre la propiedad arquimediana y el principio de buena ordenación de **N**. Para demostrar **i** supongamos que $x < y$, de modo que $x - y > 0$. Por la propiedad arquimediana de $\mathbb{R}$, existe un entero n tal que $0 < 1/n < y - x$. De nuevo, por la propiedad arquimediana (versión 2), existe un entero k tal que $k/n > x$. Por el principio de buena ordenación, existirá un k positivo mínimo con esa propiedad. Para este valor, $(k-1)/n \le x < k/n$. Si $r = k/n$, entonces

$$x < r = \frac{k-1}{n} + \frac{1}{n} \le x + \frac{1}{n} < x + (y - x) = y,$$

así que $x < r < y$, como se pedía. La parte **ii** se sigue de la parte **i** haciendo $y = x + \varepsilon$. Se deja al lector que complete la demostración. ■

1.2.18 Teorema *El intervalo unidad* $]0, 1[$ *en* $\mathbb{R}$ *no es numerable.*

Demostración Supongamos que $x_1, x_2, x_3, \ldots$ es una enumeración (lista) de los puntos de $]0, 1[$. Vamos a probar que debe existir al menos un $x \in]0, 1[$ que no aparece en la lista y, por lo tanto, no es posible establecer una correspondencia biunívoca entre N y $]0, 1[$. Expresemos cada x_k en forma decimal:

$$\begin{aligned}
x_1 &\leftrightarrow 0.a_{11}a_{12}a_{13}a_{14}a_{15}\ldots \\
x_2 &\leftrightarrow 0.a_{21}a_{22}a_{23}a_{24}a_{25}\ldots \\
x_3 &\leftrightarrow 0.a_{31}a_{32}a_{33}a_{34}a_{35}\ldots \\
&\ldots \\
x_k &\leftrightarrow 0.a_{k1}a_{k2}a_{k3}\ldots a_{kk}\ldots \\
&\ldots
\end{aligned}$$

donde cada a_{jk} es un dígito entre 0 y 9. Preferimos secuencias infinitas de 9 en lugar de expresiones decimales finitas, es decir, $1/4 = 0.2499999\ldots$. Sea $x = 0.b_1b_2b_3b_4\ldots$, donde $b_k = 5$ si $a_{kk} \ne 5$ y $b_k = 6$ si $a_{kk} = 5$; entonces $x \in]0, 1[$ pero no está en la lista porque es diferente de x_k en el k-ésimo decimal. ■

1.2.19 Corolario *Si x e y están en $\mathbb{R}$ y $x < y$, entonces el intervalo $]x, y[$ contiene un infinito numerable de números racionales y un infinito no numerable de números irracionales.*

Demostración Esto se deduce de **1.2.17** y **1.2.18**. ■

1.2.20 Proposición: La sucesión armónica *Sean $x_1 = 1$ y $x_n = 1 + 1/2 + \cdots + 1/n$, $n = 2, 3, \ldots$. Entonces x_n es monótona creciente, pero no está acotada superiormente y, por lo tanto, no converge (escribimos $x_n \to \infty$).*

Demostración Obsérvese que $x_2 = 1 + 1/2$, $x_4 = 1 + 1/2 + 1/3 + 1/4 > 1 + 1/2 + 1/4 + 1/4 = 1 + 1/2 + 1/2$, $x_8 = 1 + 1/2 + 1/3 + 1/4 + 1/5 + 1/6 + 1/7 + 1/8 > 1+1/2+1/2+4/8 = 1+1/2+1/2+1/2$. En general, por inducción, obtenemos $x_{2^n} \geq 1 + (n/2)$ y, por lo tanto, x_n no está acotada. ■

1.3.2 Proposición *Sea $S \subset \mathbb{R}$ no vacío. Entonces $b \in \mathbb{R}$ es el supremo de S sii b es una cota superior y para todo $\varepsilon > 0$ existe $x \in S$ tal que $x > b - \varepsilon$.*

Demostración Supongamos primero que $b = \sup(S)$ y que $\varepsilon > 0$. Debemos encontrar un $x \in S$ tal que $b < x + \varepsilon$. Si no existiera tal x, tendríamos $b \geq x + \varepsilon$ para todo $x \in S$; es decir, $b - \varepsilon \geq x$. Por lo tanto, $b - \varepsilon$ sería una cota superior estrictamente menor que b, con lo que b no sería el supremo, en contradicción con nuestra hipótesis.

Para la recíproca, supongamos que b satisface la condición dada. Sea d una cota superior de S. Según la definición de $\sup(S)$, debemos probar que $b \leq d$. Supongamos que $b > d$. Sea $\varepsilon = b - d$; entonces $d = b - \varepsilon$, luego $d \geq x$ para todo $x \in S$ implica que $b - \varepsilon \geq x$ o, lo que es lo mismo, que $b \geq x + \varepsilon$, de donde nuestra hipótesis no se cumple. Por lo tanto, la suposición de que $b > d$ es incorrecta y podemos concluir que $b \leq d$, como se pedía. Esto completa la prueba. ■

Nota: En esta prueba hemos usado el siguiente principio básico de la lógica: demostrar que una afirmación P implica una afirmación Q (en símbolos $P \Rightarrow Q$) equivale a demostrar que $\sim Q \Rightarrow \sim P$, donde $\sim Q$ es la negación de Q. Llamamos a $\sim Q \Rightarrow \sim P$ la ***contrapositiva*** de $P \Rightarrow Q$, mientras que $Q \Rightarrow P$ es la ***recíproca***. Aunque probar $P \Rightarrow Q$ equivale a probar $\sim Q \Rightarrow \sim P$, no es necesariamente lo mismo que probar $Q \Rightarrow P$, como lo muestran ejemplos sencillos, por ejemplo, $(x = -1)(x^2 = 1)$.

1.3.3 Proposición *Supongamos que $\varnothing \neq A \subset B \subset \mathbb{R}$. Entonces*

$$\inf A \leq \sup A \geq \inf B \leq \sup B.$$

Demostración Por ejemplo, para probar que $\inf B \leq \inf A$, sea $b = \inf B$; entonces $b \leq x$ para todo $x \in B$ y, por lo tanto, $b \leq x$ para todo $x \in A$ también, porque $A \subset B$. Entonces b es una cota inferior de A, de donde $b \leq \inf A$. ■

1.3.4 Teorema *Las siguientes afirmaciones son válidas en* $\mathbb{R}$:

i. ***Propiedad del supremo*** *Sea S un conjunto no vacío de* $\mathbb{R}$ *que tiene una cota superior; entonces S tiene supremo en* $\mathbb{R}$.

ii. ***Propiedad del ínfimo*** *Sea P un conjunto no vacío de* $\mathbb{R}$ *que tiene una cota inferior; entonces P tiene ínfimo en* $\mathbb{R}$.

Demostración

i. Vamos a dar dos demostraciones diferentes. Véase el ejercicio 46 al final del capítulo para una tercera demostración.

Método 1 Sean M una cota superior y n un entero positivo. Fijado n, considérese la sucesión $M - 1/2^n$, $M - 2/2^n$, $M - 3/2^n$, . . . que resta a M una cantidad $1/2^n$ en cada paso de la sucesión. Elijamos el primer entero k para el cual $M - k/2^n$ ya no es una cota superior. Llamemos a este entero k_n. Debe existir tal k_n porque $S \neq \varnothing$ y $\mathbb{R}$ satisface la propiedad arquimediana. Sea $b_n = M - k_n/2^n$; por construcción b_n *no* es una cota superior, pero $b_n + 1/2^n$ *sí* lo es. Como el tamaño del paso disminuye en un factor 2 cada vez que n aumenta una unidad, $b_1 \leq b_2 \leq b_3 \ldots$ Por otro lado, como $b_n \leq M$ para todo n, podemos aplicar la propiedad de completitud y asegurar que $b_n \to b$ para algún b. *También* $b \geq b_n$ (¿por qué?). Afirmamos entonces que b es el supremo. Para probar esto, sea $x \in S$. Si $x > b$, sea $\varepsilon = x - b$ y, de acuerdo con el ejemplo **1.2.14**, elijamos n tan grande que $1/2^n < \varepsilon$. Entonces

$$x = b + \varepsilon \geq b_n + \varepsilon > b_n + \frac{1}{2^n}.$$

de donde $b_n + 1/2^n$ no es una cota superior, lo que es una contradicción. De manera similar, para todo $\varepsilon > 0$ debe existir $x \in S$ tal que $b - x < \varepsilon$, por lo tanto b es el supremo, por **1.3.2**.

Método 2 Como $S \neq \varnothing$, podemos elegir $x \in S$. Escribamos $y \geq S$ si y es una cota superior de S. Para $k = 0, 1, 2, \ldots$, sea N_k el *menor* número natural para el cual $x + N_k/2^k \geq S$; N_k existe por la propiedad de buena ordenación. Sea $y_k = x + N_k/2^k$. Obsérvese que, o bien $N_{k+1} = 2 \cdot N_k$, o bien $N_{k+1} = 2 \cdot N_k - 1$, por tanto, y_k es una sucesión decreciente acotada inferiormente por x.

Al ser $\mathbb{R}$ completo, $y_k \to y$ para algún $y \in \mathbb{R}$. Probaremos que y es el supremo de S. Primero demostraremos que es una cota superior. Supongamos que $z \in S$ y $z > y$. Como $y_k \to y$, existe algún k para el cual $y_k < z$. Esto es imposible, porque cada y_k es una cota superior de S. Ahora necesitamos demostrar que y es la mínima cota superior. Por la proposición **1.3.2** es suficiente probar que para todo $\varepsilon > 0$ existe $z \in S$ tal que $y < z + \varepsilon$. Elijamos k de modo que $1/2^k < \varepsilon$, lo cual es posible por el ejemplo **1.2.14**. Por la elección de y_k, existe $z \in S$ tal que $z > y_k - 1/2^k$. Pero $y \leq y_k$, luego $z > y - 1/2^k > y - \varepsilon$, y la demostración de **i** está completa.

ii. Considérese el conjunto $-P = \{-x \mid x \in P\}$. Como P está acotado inferiormente, $-P$ está acotado superiormente, luego por **i**, $-P$ tiene una cota superior mínima $c \in \mathbb{R}$. De la definición vemos que $-c$ es la cota inferior máxima que se pide. ■

Ya hemos comentado anteriormente que las propiedades **i** y **ii** equivalen cada una al axioma de completitud para un cuerpo ordenado. Hemos probado una de las implicaciones, a saber, que el axioma de completitud implica **i** y **ii** en un cuerpo ordenado. El ejercicio 11 pide al lector que pruebe que **i** y **ii** implican el axioma de completitud.

1.4.2 Proposición *Toda sucesión convergente es una sucesión de Cauchy.*

Demostración Si $x_n \to x$, para todo $\varepsilon > 0$ existe N tal que $|x_n - x| < \varepsilon/2$ si $n \geq N$. Por lo tanto, si $n, m \geq N$, $|x_n - x_m| \leq |x_n - x| + |x - x_m| < \varepsilon/2 + \varepsilon/2 = \varepsilon$, por la desigualdad triangular, por lo que x_n es de Cauchy. ■

1.4.3 Teorema *Toda sucesión acotada en $\mathbb{R}$ tiene una subsucesión que converge a algún punto de $\mathbb{R}$.*

Demostración Supongamos que x_n es una sucesión acotada en $\mathbb{R}$, esto es, existe un entero M tal que $-M < x_n < M$ para todo n. Dividamos el intervalo $[-M, M]$ en dos mitades $[-M, 0]$ y $[0, M]$. Al menos una de ellas debe contener infinitos términos de x_n. Llamemos I_0 a dicho subintervalo y elijamos n_0 tal que $x_{n_0} \in I_0$. Ahora dividamos por la mitad I_0 y llamemos I_1 a la mitad que contenga infinitos términos de x_n. Como hay disponibles infinitos x_n, elijamos $n_1 > n_0$ tal que $x_{n_1} \in I_1$. Continuando de esta manera obtenemos sendas sucesiones de subintervalos, índices y puntos tales que

1. $I_0 \supset I_1 \supset I_2 \supset I_3 \supset \ldots.$

2. $I_k = [a_k, b_k]$ con $b_k - a_k = M/2^k$.

3. $n_0 < n_1 < n_2 < n_3 < \ldots$

4. $x_{n_k} \in I_k$. (Véase la figura 1.P-1.)

Consideremos la sucesión $a_1, a_2, a_3, \ldots$ Como $I_{k+1} \subset I_k \subset [-M, M]$, tendremos $-M \leq a_k \leq a_{k+1} \leq M$ para todo k. La sucesión es, pues, monótona creciente y acotada y, por tanto, debe converger a algún número x. Para cada k tenemos que

$$|x_{n_k} - x| \leq |x_{n_k} - a_k| + |a_k - x| \leq M/2^k + |a_k - x|.$$

Sea $\varepsilon > 0$. Por el ejemplo **1.2.14**, $1/2^k \to 0$ cuando $k \to \infty$, por tanto existe un entero K_1 tal que $1/2^k < \varepsilon/(2M)$ cuando $k \geq K_1$. Como $a_k \to x$, existe un entero K_2 tal que $|a_k - x| < \varepsilon/2$ siempre que $k \geq K_2$. Sea $K = \text{máx}(K_1, K_2)$. Si $k \geq K$, entonces $|x_{n_k} - x| < \varepsilon/2 + \varepsilon/2 = \varepsilon$; por tanto, $\text{lím}_{k\to\infty} x_{n_k} = x$, y así, la subsucesión x_{n_k} converge a x. ■

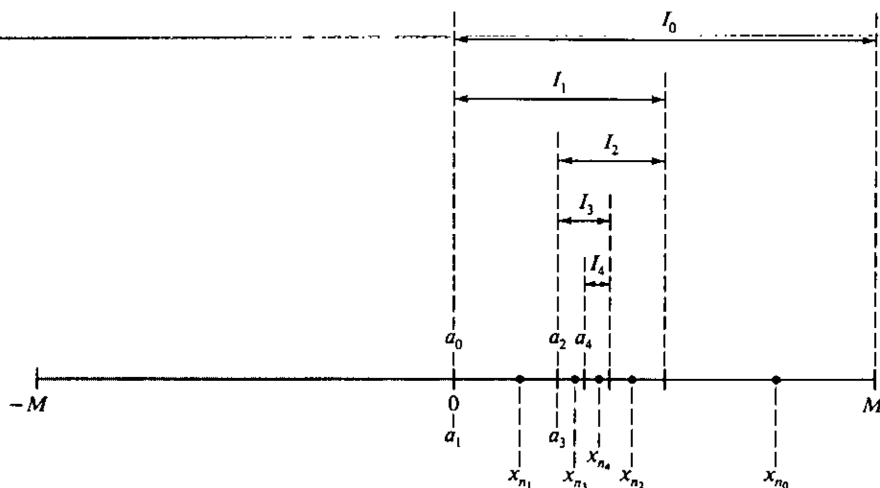


FIGURA 1.P-1 El proceso de bisección se utiliza para mostrar que toda sucesión en $[-M, M]$ tiene una subsucesión convergente

1.4.5 Corolario *Si a y b están en $\mathbb{R}$ y $a < b$, entonces toda sucesión de puntos en el intervalo $[a, b] = \{x \mid a \leq x \leq b\}$ tiene una subsucesión que converge a algún punto de $[a, b]$.*

Demostración Si x_n es una sucesión de puntos en $[a, b]$, entonces $a \leq x_n \leq b$ para todo n. La sucesión está, pues, acotada, por lo que el teorema **1.4.3** implica que existe una subsucesión $x_{n_1}, x_{n_2}, x_{n_3}, \ldots$ y un punto x en $\mathbb{R}$ tales que $\lim_{k\to\infty} x_{n_k} = x$. Por la proposición **1.2.3**, $a \leq x \leq b$, luego $x \in [a, b]$. ■

Ya vimos en el texto cómo se obtiene el teorema **1.4.4** a partir de los siguientes lemas.

1.4.6 Lema *Toda sucesión de Cauchy está acotada.*

Demostración Supongamos que x_n es una sucesión de Cauchy. Entonces hay un N tal que $|x_n - x_k| < 1$ siempre que $n \geq N$ y $k \geq N$. Por tanto, $|x_n| \leq |x_N| + 1$ para $n = N, N+1, N+2, \ldots$ Si hacemos $M = \text{máx}\{|x_1|, |x_2|, \ldots, |x_N|\} + 1$, entonces $|x_n| \leq M$ para todo n. ■

1.4.7 Lema *Si una subsucesión de Cauchy converge a x, entonces la sucesión misma converge a x.*

Demostración Sea x_n una sucesión de Cauchy que tiene una subsucesión x_{n_k} que converge a x. Sea $\varepsilon > 0$, entonces existe N tal que $|x_n - x_m| < \varepsilon/2$ cuando $n \geq N$ y $m \geq N$. Como x_{n_k} es una subsucesión, existe $n_0 > N$ tal que $|x_{n_0} - x| < \varepsilon/2$. Si $m > N$, entonces

$$|x_m - x| = |x_m - x_{n_0} + x_{n_0} - x| \leq |x_m - x_{n_0}| + |x_{n_0} - x| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Por tanto, x_n converge a x. ■

1.5.2 Proposición *Sea x_n una sucesión en $\mathbb{R}$ y sea $x \in \mathbb{R}$.*

i. *x es un punto límite de x_n si y sólo si para todo $\varepsilon > 0$ y para todo N, existe un índice $n > N$ tal que $|x_n - x| < \varepsilon$.*

ii. *x es un punto límite de x_n si y sólo si existe una subsucesión de x_n que converja a x.*

iii. *$x_n \to x$ si y sólo si toda subsucesión de x_n converge a x.*

iv. *$x_n \to x$ si y sólo si toda subsucesión está acotada y x es su único punto límite.*

v. *$x_n \to x$ si y sólo si toda subsucesión de x_n tiene a su vez una subsucesión que converge a x.*

Demostración

i. Si x es un punto límite de x_n y $\varepsilon > 0$, entonces hay infinitos índices para los cuales $|x_n - x| < \varepsilon$. Sólo un número finito de ellos es menor que N, luego entonces uno de ellos debe ser mayor que n. Para la recíproca, elegimos los índices de forma inductiva: escogemos n_1 tal que $|x_{n_1} - x| < \varepsilon$, después $n_2 > n_1$ tal que $|x_{n_2} - x| < \varepsilon$, después $n_3 > n_2$ tal que $|x_{n_3} - x| < \varepsilon$, y así sucesivamente. Obtenemos $n_1 < n_2 < n_3 < \cdots$ tales que $|x_{n_k} - x| < \varepsilon$ y, por lo tanto, x es un punto límite.

ii. Para probar esta parte refinaremos el proceso de selección que acabamos de hacer. Si x es un punto límite, seleccionemos el índice n_1 de modo que $|x_{n_1} - x| < 1$. Usando **i**, elijamos $n_2 > n_1$ tal que $|x_{n_2} - x| < 1/2$. Continuemos de esta forma, utilizando **i** reiteradamente, y elijamos los índices $n_1 < n_2 < n_3 < n_4 < \cdots$ de modo que $|x_{n_k} - x| < 1/k$. Esto produce una subsucesión que converge a x, ya que si $\varepsilon > 0$ sabemos que existe un entero K tal que $1/K < \varepsilon$; entonces, $k \geq K$ implica $|x_{n_k} - x| < 1/k \leq 1/K < \varepsilon$. Para la recíproca, si hay una subsucesión convergente a x y $\varepsilon > 0$, entonces hay un índice tal que los índices posteriores a éste en la subsucesión distan menos que ε de x. Por lo tanto x es un punto límite.

iii. Si $x_n \to x$ y x_{n_k} es una subsucesión, sea $\varepsilon > 0$ y elijamos N de tal modo que $|x_n - x| < \varepsilon$ para $n \geq N$. Como $n_1 < n_2 < n_3 < \cdots$, por inducción se demuestra que $n_k \geq k$. Por lo tanto $k \geq N$ implica $n_k \geq N$, y, en consecuencia, $|x_{n_k} - x| < \varepsilon$; la subsucesión converge a x. Para demostrar la recíproca, probaremos su contrapositiva, es decir, si x_n fuera una sucesión que no convergiera a x, entonces por **i** existiría $\varepsilon > 0$ tal que para cada N habría un índice mayor que N tal que $|x_n - x| > \varepsilon$. Usando este hecho repetidamente podemos seleccionar $n_1 < n_2 < n_3 < \cdots$ tales que $|x_{n_k} - x| > \varepsilon$. Esto nos da una subsucesión que no converge a x.

iv. Si $x_n \to x$, entonces sabemos que la sucesión está acotada y que toda subsucesión también converge a x; por lo tanto, x es el único punto límite. Para la recíproca, supongamos que la sucesión está acotada y que x es su único punto límite. Si la sucesión no converge a x deberán existir un $\varepsilon > 0$ y una subsucesión $x_{k(j)}$ tales que $|x - x_{k(j)}| \geq \varepsilon$ para todo j. Como dicha subsucesión está acotada, debe tener una sub-subsucesión convergente. El límite de esta sub-subsucesión es un punto límite de la sucesión original diferente de x; pero no existen tales puntos, por lo que $x_k \to x$.

v. Si $x_n \to x$, entonces, por **iii**, toda subsucesión converge a x, por lo cual no necesitamos pasar a una sub-subsucesión. Si toda subsucesión tiene una sub-subsucesión convergente a x, entonces x es un punto límite de toda subsucesión y por lo tanto de x_n. Si existiera otro punto límite $y \neq x$, entonces, por **ii**, habría una subsucesión convergente a él. Por **iv**, y sería el único punto límite de esa subsucesión, y, en consecuencia, no podría tener una sub-subsucesión convergente a x. Así pues, x debe ser el punto límite de x_n, con lo cual, por **iv**, $x_n \to x$. ■

1.5.5 Proposición *Sea x_n una sucesión en $\mathbb{R}$. Entonces*

i. *Si x_n está acotada inferiormente, un número a es igual al* lím inf x_n *si y sólo si*

a. *Para todo $\varepsilon > 0$ existe un N tal que $a - \varepsilon < x_n$ siempre que $n \leq N$, y*

b. *Para todo $\varepsilon > 0$ y para todo M existe un $n > M$ tal que $x_n < a + \varepsilon$.*

ii. *Si x_n está acotada superiormente, un número b es igual al* lím sup x_n *si y sólo si*

a. *Para todo $\varepsilon > 0$ existe un N tal que $x_n < b + \varepsilon$ siempre que $n \geq N$, y*

b. *Para todo $\varepsilon > 0$ y para todo M existe un $n > M$ tal que $b - \varepsilon < x_n$.*

Demostración Supongamos que $a = \text{lím inf}\, x_n$. Entonces $a = \inf\{x \mid x \text{ es un punto límite de } x_n\}$, por definición. Así pues, dado $\varepsilon > 0$, existe un punto límite x tal que $x - \varepsilon/2 < a$; además, $a \leq x$ para todos los puntos límite. Si x es un punto límite, entonces hay infinitos x_n para los cuales $|x_n - x| < \varepsilon/2$ y, por lo tanto, tales x_n satisfacen $x_n - \varepsilon < a$, lo que prueba **ib**. Si **ia** es falsa, entonces existen $\varepsilon > 0$ e infinitos x_n que satisfacen $x_n < a - \varepsilon$. De hecho, elijamos una subsucesión x_{n_k} tal que $x_{n_k} < a - \varepsilon$; como x_n está acotada inferiormente, x_{n_k} está acotada y, por lo tanto, tiene, a su vez, una subsucesión que converge, digamos, a x. Entonces $x \leq a - \varepsilon$, de donde x es un punto límite menor que a, lo cual es una contradicción. La recíproca se prueba de la misma manera, y la demostración de **ii** es similar. ∎

1.5.6 Proposición *Sea x_n una sucesión en $\mathbb{R}$. Entonces*

$$\text{lím sup}\, x_n = \inf\{\sup\{x_{n+1}, x_{n+2}, x_{n+3}, \ldots\} \mid n = 1, 2, 3, \ldots\}$$
$$\text{lím inf}\, x_n = \sup\{\inf\{x_{n+1}, x_{n+2}, x_{n+3}, \ldots\} \mid n = 1, 2, 3, \ldots\}.$$

Demostración Probemos la igualdad para lím sup x_n. En primer lugar, podemos suponer que x_n está acotada superiormente, porque de otra forma ambos miembros serían $+\infty$ (entendiendo que $\inf\{\infty\} = \infty$). Sea $b = \text{lím sup}\, x_n$ y $b_n = \sup\{x_{n+1}, x_{n+2}, \ldots\}$, como en el texto, de manera que $b_1 \geq b_2 \geq \cdots$ y $b_n \to \inf\{b_1, b_2, \ldots\}$ (si b_n no estuviera acotada inferiormente, ambos miembros serían $-\infty$, un caso que dejamos al lector). Necesitamos demostrar que $b_n \to b$. Para todo $\varepsilon > 0$ existe N tal que $x_n < b + \varepsilon$ siempre que $n \geq N$, por **1.5.5ib;** por lo tanto, $b_n \leq b + \varepsilon$ para $n \geq N$. Por otro lado, como $b - \varepsilon < x_n$, para $n \geq M$ arbitrariamente grande, tendremos entonces que $b - \varepsilon \leq b_n$ cuando $n \geq M$. Así pues, para n grande, $b - \varepsilon \leq b_n \leq b + \varepsilon$, luego $b_n \to b$. ∎

1.5.7 Proposición *Sea x_n una sucesión dada en $\mathbb{R}$. Entonces*

i. $\lim\inf x_n \leq \lim\sup x_n$.

ii. *Si* $x_n \leq M$ *para todo n, entonces* $\lim\sup x_n \leq M$.

iii. *Si* $M \leq x_n$ *para todo n, entonces* $\lim\inf x_n \geq M$.

iv. $\lim\sup x_n = +\infty$ *si y sólo si x_n no está acotada superiormente.*

v. $\lim\inf x_n = -\infty$ *si y sólo si x_n no está acotada inferiormente.*

vi. *Si x es un punto límite de x_n, entonces* $\lim\inf x_n \leq x \leq \lim\sup x_n$.

vii. *Si* $a = \lim\inf x_n$ *es finito, entonces es un punto límite de x_n.*

viii. *Si* $b = \lim\sup x_n$ *es finito, entonces es un punto límite de x_n.*

ix. $x_n \to x \in \mathbb{R}$ *si y sólo si* $\lim\sup x_n = \lim\inf x_n = x \in \mathbb{R}$.

Demostración

i. Dada una sucesión x_n en $\mathbb{R}$, sea $S_n = \{x_{n+1}, x_{n+2}, \ldots\}$, $a_n = \inf S_n$ (definamos $a_n = -\infty$ si S_n no está acotado inferiormente) y $b_n = \sup S_n$ (definamos $b_n = \infty$ si S_n no está acotado superiormente). De los comentarios que preceden al enunciado de la proposición **1.5.6**, cada b_k es una cota superior de $\{a_1, a_2, a_3, \ldots\}$, por lo tanto, $\lim\inf x_n = \sup\{a_1, a_2, a_3, \ldots\} \leq b_k$ para todo k. Esto es válido para *todos* los valores de k, luego $\lim\inf x_n \leq \inf\{b_1, b_2, b_3, \ldots\} = \lim\sup x_n$.

ii. Como $x_n \leq M$ para todo n, $\sup S_n \leq M$ para todo n, y, por lo tanto, $\lim\sup x_n = \inf(\sup S_n) \leq M$. La demostración de **iii** es similar.

iv. Si x_n no está acotada superiormente, $\lim\sup x_n = +\infty$, por definición. La recíproca se sigue de **ii.** La demostración de **v** es similar.

vi. Supongamos que x es un punto límite de x_n. Entonces para todo n existe un índice $k > n$ para el cual $x - \varepsilon < x_k < x + \varepsilon$, esto es, $S_n \cap\,]x - \varepsilon, x + \varepsilon[\, \neq \varnothing$. Por lo tanto, para todo n,

$$x - \varepsilon \leq \sup S_n \text{ e } \inf S_n \leq x + \varepsilon.$$

Como esto se cumple para todo $\varepsilon > 0$, tendremos

$$\inf S_n \leq x \leq \sup S_n \text{ para todo } n,$$

luego

$$\lim\inf x_n = \sup_n(\inf S_n) \leq x \leq \inf_n(\sup S_n) = \lim\sup x_n.$$

vii. Por definición, $a = \sup\{a_1, a_2, a_3, \ldots\}$, donde $a_n = \inf\{x_{n+1}, x_{n+2}, \ldots\}$. Claramente, $a_1 \le a_2 \le a_3 \le \cdots \le a$. Como a es finito e igual al supremo de los a_k, esta sucesión debe converger hacia a. Sean ε y N dados. Existe un N_0 tal que $a - \varepsilon < a_n \le a$ siempre que $n \ge N_0$. Eligiendo $n > \text{máx}(N, N_0)$ tenemos $a - \varepsilon < \inf S_n \le a$. Por lo tanto, S_n debe intersecar el intervalo $]a - \varepsilon, a + \varepsilon[$, esto es, existe un índice $k > n \ge N$ tal que $|x_k - a| < \varepsilon$, y, por lo tanto, a es un punto límite.

viii. La demostración se deja como ejercicio.

ix. Si $x_n \to x$, entonces para todo $\varepsilon > 0$ existe un N tal que $x - \varepsilon < x_n < x + \varepsilon$ para todo $n \ge N$, de donde, por **1.5.5i**, $\text{lím inf}\, x_n = x$ y por **1.5.5ii**, $\text{lím sup}\, x_n = x$.

Recíprocamente, si se cumplen las condiciones de **1.5.5** con $x = a = b$, sea N el mayor de los N correspondientes a **1.5.5ia** y **1.5.5iia**, de tal manera que $x - \varepsilon < x_n < x + \varepsilon$ si $n \ge N$ y, por tanto, $x_n \to x$ cuando $n \to \infty$. ■

1.6.6 Teorema *Para vectores de $\mathbb{R}^n$, tenemos*

I. *Propiedades del producto interno:*

i. $\langle x, x\rangle \ge 0$. *positividad*

ii. $\langle x, x\rangle = 0$ *sii* $x = 0$. *no degeneración*

iii. $\langle x, y + w\rangle = \langle x, y\rangle + \langle x, w\rangle$. *distributividad*

iv. $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$ *para* $\alpha \in \mathbb{R}$. *multiplicatividad*

v. $\langle x, y\rangle = \langle y, x\rangle$. *simetría*

II. *Propiedades de la norma:*

i. $\|x\| \ge 0$. *positividad*

ii. $\|x\| = 0$ *sii* $x = 0$. *no degeneración*

iii. $\|\alpha x\| = |\alpha|\, \|x\|$ *para todo* $\alpha \in \mathbb{R}$. *multiplicatividad*

iv. $\|x + y\| \le \|x\| + \|y\|$. *desigualdad triangular*

III. *Propiedades de la distancia:*

i. $d(x, y) \ge 0$. *positividad*

ii. $d(x, y) = 0$ *sii* $x = y$. *no degeneración*

iii. $d(x, y) = d(y, x)$. *simetría*

iv. $d(x, y) \le d(x, z) + d(z, y)$. *desigualdad triangular*

IV. *Desigualdad de Cauchy-Schwarz:*

$$|\langle x, y\rangle| \leq \|x\|\ \|y\|.$$

Demostración

I. Todas las propiedades son inmediatas excepto **IV**, que se probará en **1.7.9**.

II y **III** se probarán como consecuencias generales de **I** en **1.7.5**, **1.7.9**, y **1.7.10**. ■

1.6.7 Proposición *Si $v = (v_1, \ldots, v_n)$ y $w = (w_1, \ldots, w_n)$ son vectores de $\mathbb{R}^n$ y definimos $\rho(v, w)$, = $\max\{|v_1 - w_1|, |v_2 - w_2|, \ldots, |v_n - w_n|\}$, entonces*

$$\rho(v, w) \leq \|v - w\| \leq \sqrt{n}\rho(v, w).$$

Demostración Es fácil ver que

$$|v_i - w_i| = \sqrt{|v_i - w_i|^2} \leq \sqrt{\sum_{j=1}^{n} |v_j - w_j|^2} = \|v - w\|.$$

También

$$\|v - w\| = \sqrt{\sum_{j=1}^{n}\{|v_j - w_j|^2\}} \leq \sqrt{\sum_{j=1}^{n}\{\max_j |v_j - w_j|^2\}}$$
$$= \sqrt{n\rho(v, w)^2} = \sqrt{n}\rho(v, w),$$

lo cual demuestra el resultado. ■

1.7.5 Proposición *Si $(V, \|\cdot\|)$ es un espacio vectorial normado y $d(v, w)$ está definida por*

$$d(v, w) = \|v - w\|,$$

entonces d es una métrica sobre V.

Demostración

i. $d(v,w) \geq 0$, porque $d(v,w) = ||v - w|| \geq 0$.

ii. Para probar $d(v,w) = 0 \Leftrightarrow v = w$, obsérvese que $d(v,w) = 0 \Leftrightarrow ||v - w|| = 0 \Leftrightarrow v - w = 0 \Leftrightarrow v = w$.

iii. $d(v,w) = d(w,v)$, porque $d(v,w) = ||v-w|| = ||-(w-v)|| = |-1|\ ||w-v|| = ||w - v|| = d(w - v)$

iv. Por la desigualdad triangular, $d(x,y) = ||x - y|| = ||(x - z) + (z - y)|| \leq ||x - z|| + ||z - y|| = d(x,z) + d(z,y)$. ■

1.7.9 Desigualdad de Cauchy-Schwarz *Si $(V, \langle\cdot,\cdot\rangle)$ es un espacio con producto interno, entonces*

$$|\langle x,y\rangle| \leq \sqrt{\langle x,x\rangle}\ \sqrt{\langle y,y\rangle}$$

para todo x e y en V.

Demostración Si x o y son 0, entonces $\langle x, y\rangle = 0$, y, por lo tanto, la desigualdad se cumple. Así pues, podemos suponer que $x \neq 0$ e $y \neq 0$. Entonces $\langle x, x\rangle > 0$ y $\langle y, y\rangle > 0$. Para todo $\alpha \in \mathbb{R}$, tenemos

$$0 \leq \langle \alpha x + y, \alpha x + y\rangle = \alpha^2\langle x,x\rangle + 2\langle x,y\rangle\alpha + \langle y,y\rangle$$

usando las propiedades básicas del producto interno. Si definimos $\langle x, x\rangle = a$, $2\langle x, y\rangle = b$ y $\langle y, y\rangle = c$, esta desigualdad se convierte en

$$a\alpha^2 + b\alpha + c \geq 0 \quad \text{para todo } \alpha \in \mathbb{R}.$$

Sabemos (ya que $a > 0$) que esta expresión cuadrática tiene un mínimo para $\alpha = -b/2a$. Si sustituimos por este valor obtenemos

$$a\left(\frac{b^2}{4a^2}\right) + b\left(-\frac{b}{2a}\right) + c \geq 0; \quad \text{es decir,}\ c \geq \frac{b^2}{4a}.$$

Como $a > 0$, esto significa que

$$b^2 \leq 4ac;\ \text{es decir,}\ |\langle x,y\rangle|^2 \leq \langle x,x\rangle\langle y,y\rangle.$$

Tomando la raíz cuadrada obtenemos lo que queríamos. ■

1.7.10 Proposición *Si $(V, \langle\cdot,\cdot\rangle)$ es un espacio con producto interno y $\|\cdot\|$ está definida para cada v en V por*

$$||v|| = \sqrt{\langle v,v\rangle},$$

entonces $\|\cdot\|$ es una norma sobre V.

Demostración La única propiedad no trivial es la desigualdad triangular, la cual se probó en el texto. ■

Ejemplos resueltos del capítulo 1

Ejemplo 1.1 *Pruébese que, para los números reales,*

a. $x \leq |x|,\ -|x| \leq x$.

b. $|x| \leq a$ *sii* $-a \leq x \leq a$, *donde* $a \geq 0$.

c. $|x+y| \leq |x| + |y|$.

Solución

a. Si $x \geq 0$, entonces $|x| = x$; si $x < 0$, entonces $|x| \geq x$, porque $x \geq 0$. En cualquier caso $x \leq |x|$. La otra afirmación es similar.

b. Si $x \geq 0$, entonces debemos demostrar que $0 \leq x \leq a$ sii $-a \leq x \leq a$, lo cual es evidente. De manera similar, si $x < 0$ la afirmación toma la forma $0 \leq -x \leq a$ sii $-a \leq x \leq a$, lo que también es evidente. En este caso hemos utilizado el hecho de que si $c \leq 0$, entonces $0 \leq x \leq y$ sii $0 \geq cx \leq cy$.

c. Por **a**, $-|x| \leq x \leq |x|$ y $-|y| \leq y \leq |y|$. Sumamos y obtenemos $-(|x| + |y|) \leq x + y \leq |x| + |y|$. Entonces, por **b**, $|x+y| \leq |x| + |y|$. Esto también se puede probar por casos, como se indica en el texto. Además, obsérvese que éste es un caso especial de la desigualdad triangular en $\mathbb{R}^n$; véase el teorema **1.6.6IIiv.** ♦

Ejemplo 1.2 *Sea S un conjunto no vacío de $\mathbb{R}$ acotado superiormente y sea $x =$* sup(S). *Demuéstrese que existe una sucesión $x_1, x_2, \ldots$ en S tal que $x_k \to x$.*

Solución Para cada k, usamos la proposición **1.3.2** con objeto de encontrar $x_k \in S$ tal que $x_k \leq x < x_k + 1/k$. Entonces $x_k \to x$, porque para cada $\varepsilon > 0$ dado, elegimos $N \geq 1/\varepsilon$; entonces $k \geq N$ implica $x_k \leq x < x_k + \varepsilon$, es decir, $|x - x_k| < \varepsilon$. ♦

Ejemplo 1.3 *Para los números $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n$ y $z_1, z_2, \ldots, z_n$ pruébese que*

$$\left(\sum_{i=1}^{n} x_i y_i z_i\right)^4 \leq \left(\sum_{i=1}^{n} x_i^4\right)\left(\sum_{i=1}^{n} y_i^2\right)^2\left(\sum_{i=1}^{n} z_i^4\right)$$

Solución La desigualdad de Cauchy-Schwarz (teorema **1.6.6IV**) dice que

$$\left(\sum w_i y_i\right)^2 \le \left(\sum w_i^2\right)\left(\sum y_i^2\right).$$

Aplicando esta desigualdad a los números $w_i = x_i z_i$ e y_i se obtiene

$$\left(\sum x_i y_i z_i\right)^2 \le \left(\sum (x_i z_i)^2\right)\left(\sum y_i^2\right).$$

Aplicándolo de nuevo a x_i^2 y z_i^2 se sigue que

$$\left(\sum x_i^2, z_i^2\right)^2 \le \left(\sum x_i^4\right)\left(\sum z_i^4\right) \text{ es decir, } \left(\sum (x_i z_i)^2\right) \le \left(\sum x_i^4\right)^{1/2}\left(\sum z_i^4\right)^{1/2}$$

luego

$$\left(\sum x_i y_i z_i\right)^2 \le \left(\sum x_i^4\right)^{1/2}\left(\sum z_i^4\right)^{1/2}\left(\sum y_i^2\right).$$

Elevando al cuadrado ambos miembros obtenemos el resultado. (Hemos utilizado el hecho de que si $a, b \ge 0$, entonces $a \le b$ sii $a^2 \le b^2$.) ♦

Ejemplo 1.4 *Supongamos que $x \in \mathbb{R}$ y $x > 0$; pruébese que existe un número irracional entre 0 y x.*

Solución Si x es racional, entonces, como $\sqrt{2}$ es irracional (ejercicio 2), también lo es $x/\sqrt{2}$ (¿por qué?) y está entre 0 y x. Por otro lado, si x es irracional, entonces $x/2$ es irracional (¿por qué?) y está entre 0 y x. ♦

Ejemplo 1.5 *Sean A y B dos conjuntos no vacíos de $\mathbb{R}$ que están acotados superiormente. Sean $a = \sup(A)$ y $b = \sup(B)$ y definamos el conjunto C mediante $C = \{xy \mid x \in A, y \in B\}$. Demuéstrese que, en general, $ab \neq \sup(C)$. Si $a < 0$ y $b < 0$, pruébese que $ab = \inf(C)$. Si $a > 0$, $b > 0$ y todos los elementos de A y B son positivos, pruébese que $ab = \sup(C)$.*

Solución Como contraejemplo, sea $A = \{x \in \mathbb{R} \mid -10 < x < -1\} =]-10, -1[$ y $B =]0, 1/2[$, así que $a = -1$, $b = 1/2$, y $ab = -1/2$. Pero $C =]-5, 0[$ y $\sup(C) = 0$.

Ahora probemos que si $a < 0$ y $b < 0$, entonces $ab = \inf(C)$. Para esto usamos el análogo a la proposición **1.3.2** para supremos. Primero, sean $x \in A$ e $y \in B$; queremos probar que $xy \ge ab$. Pero $x \le a$ e $y \le b$, o bien $-x \ge -a \ge 0$ y $-y \ge -b \ge 0$, y, por lo tanto (aplicando a $\mathbb{R}$ el axioma **III.16** para cuerpos ordenados), $(-x)(-y) \ge (-a)(-b)$, es

decir, $xy \geq ab$. Dado $\varepsilon > 0$, buscamos $x \in A$ e $y \in B$ de tal manera que $ab > xy - \varepsilon$, o, de otro modo, $|ab - xy| < \varepsilon$. Elijamos x e y de tal modo que $a < x + \varepsilon/[2(|b| + 1)]$, $b < y + \varepsilon/[2|a|]$ y $b < y + 1$. Entonces, dado que $|uv| = |u|\,|v|$ y $|y| < |b| + 1$, obtenemos

$$\begin{aligned} |ab - xy| &\leq |ab - ay| + |ay - xy| \\ &= |a|\,|b - y| + |a - x|\,|y| < |a|\frac{\varepsilon}{2|a|} + \frac{\varepsilon}{2(|b| + 1)}(|b| + 1) = \varepsilon \end{aligned}$$

(usando la desigualdad triangular).

La última afirmación se puede probar de manera similar. ♦

Ejercicios del capítulo 1

1. Para cada uno de los siguientes conjuntos S, encuéntrese $\sup(S)$ e $\inf(S)$, si es que existen:

a. $\{x \in \mathbb{R} \mid x^2 < 5\}$

b. $\{x \in \mathbb{R} \mid x^2 > 7\}$

c. $\{1/n \mid n \text{ es entero y } n > 0\}$

d. $\{-1/n \mid n \text{ es entero y } n > 0\}$

e. $\{.3, .33, .333, \ldots\}$

2. Repásese la demostración de que $\sqrt{2}$ es irracional y generalícese el resultado para $\sqrt{k}$, con k un entero positivo que no sea un cuadrado perfecto.

3. **a.** Sea $x \geq 0$ un número real tal que para cualquier $\varepsilon > 0$, $x \leq \varepsilon$. Pruébese que $x = 0$.

b. Sea $S =]0, 1[$. Pruébese que para cualquier $\varepsilon > 0$ existe un $x \in S$ tal que $x < \varepsilon$.

4. Pruébese que $d = \inf(S)$ sii d es una cota inferior de S y para cualquier $\varepsilon > 0$ existe un $x \in S$ tal que $d \geq x - \varepsilon$.

5. Sea x_n una sucesión monótona creciente acotada superiormente y considérese el conjunto $S = \{x_1, x_2, \ldots\}$. Pruébese que x_n converge al $\sup(S)$. Enúnciese un resultado similar para sucesiones decrecientes.

6. Sean A y B dos conjuntos no vacíos de números reales con la propiedad de que $x \leq y$ para todo $x \in A$ e $y \in B$. Pruébese que existe un número $c \in \mathbb{R}$ tal que $x \leq c \leq y$ para todo $x \in A$ e $y \in B$. Encuéntrese un contraejemplo a este enunciado para números racionales (esto, de hecho, equivale al axioma de completitud y es la base para una formulación alternativa de dicho axioma, conocida como **cortaduras de Dedekind**).

7. Para conjuntos no vacíos A, $B \subset \mathbb{R}$, sea $A + B = \{x + y \mid x \in A \text{ e } y \in B\}$. Pruébese que $\sup(A + B) = \sup(A) + \sup(B)$.

8. Para conjuntos no vacíos A, $B \subset \mathbb{R}$, determínese cuáles de los siguientes enunciados son verdaderos. Demuéstrense los enunciados verdaderos y búsquese un contraejemplo para los falsos:

a. $\sup(A \cap B) \leq \inf\{\sup(A), \sup(B)\}$.

b. $\sup(A \cap B) = \inf\{\sup(A), \sup(B)\}$.

c. $\sup(A \cup B) \geq \sup\{\sup(A), \sup(B)\}$.

d. $\sup(A \cup B) = \sup\{\sup(A), \sup(B)\}$.

9. Sea x_n una sucesión acotada de números reales e $y_n = (-1)^n x_n$. Pruébese que lím sup $y_n \leq$ lím sup $|x_n|$. ¿Son, de hecho, iguales? Escríbase una desigualdad similar para el lím inf.

10. Verifíquese que la métrica acotada del ejemplo **1.7.2d** es, en efecto, una métrica.

11. Pruébese que tanto la proposición **i** como la **ii** del teorema **1.3.4** implican el axioma de completitud para un cuerpo ordenado.

12. En un espacio con producto interno, pruébese que

a. $2\|x\|^2 + 2\|y\|^2 = \|x + y\|^2 + \|x - y\|^2$ (***ley del paralelogramo***).

b. $\|x + y\| \, \|x - y\| \leq \|x\|^2 + \|y\|^2$.

c. $4\langle x, y\rangle = \|x + y\|^2 - \|x - y\|^2$ (***identidad de polarización***).

Interprétense los resultados geométricos en téminos del paralelogramo formado por x e y.

13. ¿Cuál es el complemento ortogonal en $\mathbb{R}^4$ del espacio generado por $(1, 0, 1, 1)$ y $(-1, 2, 0, 0)$?

14. **a.** Demuéstrese la ***identidad de Lagrange***

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 = \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right) - \sum_{1 \leq i < j \leq n} (x_i y_j - x_j y_i)^2$$

y utilícese para dar otra demostración de la desigualdad de Cauchy-Schwarz.

b. Pruébese que

$$\left(\sum_{i=1}^{n} (x_i + y_i)^2\right)^{1/2} \leq \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} + \left(\sum_{i=1}^{n} y_i^2\right)^{1/2}$$

15. Sea x_n una sucesión en $\mathbb{R}$ tal que $d(x_n, x_{n+1}) \leq d(x_{n-1}, x_n)/2$. Pruébese que x_n es una sucesión de Cauchy.

16. Pruébese el teorema **1.6.4**. De hecho, dados los espacios vectoriales $V_1, \ldots, V_n$, pruébese que $V = V_1 \times \cdots \times V_n$ es un espacio vectorial.

17. Sea $S \subset \mathbb{R}$ acotado inferiormente y no vacío. Pruébese que $\inf(S) = \sup\{x \in \mathbb{R} \mid x$ es una cota inferior de $S\}$.

18. Pruébese que en $\mathbb{R}$, $x_n \to x$ sii $-x_n \to -x$. En consecuencia, pruébese que el axioma de completitud es equivalente al enunciado que dice que toda sucesión decreciente $x_1 \geq x_2 \geq x_3 \cdots$ acotada inferiormente converge. Demuéstrese que el límite de la sucesión es $\inf\{x_1, x_2, \ldots\}$.

19. Expresemos $x = (1, 1, 1) \in \mathbb{R}^3$ en la forma $x = y_1 f_1 + y_2 f_2 + y_3 f_3$, donde $f_1 = (1, 0, 1)$, $f_2 = (0, 1, 1)$ y $f_3 = (1, 1, 0)$. Calcúlense las componentes y_i.

20. Sean S y T subespacios ortogonales de $\mathbb{R}^n$ diferentes de cero. Demuéstrese que si S y T son complementos ortogonales (es decir, S y T generan todo $\mathbb{R}^n$), entonces $S \cap T = \{0\}$ y $\dim(S) + \dim(T) = n$, donde $\dim(S)$ denota la dimensión de S. Dense ejemplos en $\mathbb{R}^3$ en los que la condición $\dim(s) + \dim(T) = n$ se cumpla y ejemplos en los que no. ¿Puede no cumplirse en $\mathbb{R}^2$?

21. Pruébese que la sucesión del ejemplo resuelto **1.2** se puede elegir de manera que sea creciente.

22. **a.** Si x_n e y_n son sucesiones acotadas en $\mathbb{R}$, pruébese que

$$\limsup(x_n + y_n) = \limsup x_n + \limsup y_n.$$

b. ¿Es cierta la regla del producto para los límites superiores?

23. Sea $P \subset \mathbb{R}$ un conjunto tal que $x \geq 0$ para todo $x \in P$ y tal que para todo entero k existe un $x_k \in P$ que cumple $kx_k \leq 1$. Pruébese que $0 = \inf(P)$.

24. Si $\sup(P) = \sup(Q)$ e $\inf(P) = \inf(Q)$, ¿será $P = Q$?

25. Decimos que $P \leq Q$ si para todo $x \in P$ existe $y \in Q$ tal que $x \leq y$.

a. Si $P \leq Q$, entonces pruébese que $\sup(P) \leq \sup(Q)$.

b. ¿Es cierto que $\inf(P) \leq \inf(Q)$?

c. Si $P \leq Q$ y $Q \leq P$, ¿es $P = Q$?

26. Supongamos que $A = \{a_{m,n} \mid m = 1, 2, 3, \ldots$ y $n = 1, 2, 3, \ldots\}$ es un conjunto acotado y que $a_{m,n} \geq a_{p,q}$ siempre que $m \geq p$ y $n \geq q$. Pruébese que

$$\lim_{n \to \infty} a_{n,n} = \sup A.$$

27. Sean $S = \{(x, y) \in \mathbb{R}^2 \mid xy > 1\}$ y $B = \{d((x, y), (0, 0)) \mid (x, y) \in S\}$. Hállese $\inf(B)$.

28. Sea x_n una sucesión convergente en $\mathbb{R}$ y definamos $A_n = \sup\{x_n, x_{n+1}, \ldots\}$ y $B_n = \inf\{x_n, x_{n+1}, \ldots\}$. Pruébese que A_n converge al mismo límite que B_n, el cual es el mismo que el límite de x_n.

29. Demuéstrese que para cualquier $x \in \mathbb{R}$ que verifique $x \geq 0$, existe $y \in \mathbb{R}$ tal que $y^2 = x$.

30. Sea V el espacio vectorial $C([0, 1])$ con la norma $\|f\|_\infty = \sup\{|f(x)| \mid x \in [0, 1]\}$. Pruébese que la ley del paralelogramo no es válida y conclúyase que esta norma no proviene de ningún producto interno. (Véase el ejercicio 12.)

31. Sean $A, B \subset \mathbb{R}$ y $f : A \times B \to \mathbb{R}$ acotada. ¿Es cierto que

$$\sup\{f(x,y) \mid (x,y) \in A \times B\} = \sup\{\sup\{f(x,y) \mid x \in A\} \mid y \in B\}$$

o lo que es lo mismo, usando una notación diferente,

$$\sup_{(x,y)\in A\times B} f(x,y) = \sup_{y\in B}\left(\sup_{x\in A} f(x,y)\right)?$$

32. **a.** Dése una definición aceptable que exprese el significado de $\lim_{n\to\infty} x_n = \infty$.

b. Sea $x_1 = 1$ y definamos inductivamente $x_{n+1} = (x_1 + \cdots + x_n)/2$. Demuéstrese que $x_n \to \infty$.

33. **a.** Pruébese que $(\log x)/x \to 0$ cuando $x \to \infty$. (Puede consultarse cualquier texto de cálculo y utilizar, por ejemplo, la regla de l'Hôpital.)

b. Pruébese que $n^{1/n} \to 1$ cuando $n \to \infty$.

Los ejercicios 34–45 tratan de números complejos.

34. Exprésense los siguientes números complejos en la forma $a + bi$:

a. $(2 + 3i)(4 + i)$

b. $(8 + 6i)^2$

c. $(1 + 3/(1 + i))^2$

35. ¿Cuál es el complejo conjugado del número $(3 + 8i)^4/(1 + i)^{10}$?

36. Encuéntrense las soluciones de:

a. $(z + 1)^2 = 3 + 4i$.

b. $z^4 - i = 0$.

37. Encuéntrense las soluciones de $z^2 = 3 - 4i$.

38. Si a es real y z es complejo, pruébese que $\text{Re}(az) = a \text{ Re } z$ y que $\text{Im}(az) = a \text{ Im } z$. Generalícese probando que Re: $\mathbb{C} \to \mathbb{R}$ es una función lineal real, es decir, que Re $(az + bw) = a \text{ Re } z + b \text{ Re } w$ para a, b reales y z, w complejos.

39. Hállense las partes reales e imaginarias de los siguientes complejos, donde $z = x + iy$:

a. $1/z^2$

b. $1/(3z + 2)$

40. **a.** Fijemos el número complejo $z = x + iy$ y consideremos la función lineal φ_z: $\mathbb{R}^2 \to \mathbb{R}^2$ (es decir, de $\mathbb{C} \to \mathbb{C}$) definida por $\varphi_z(w) = z \cdot w$ (esto es, la multiplicación por z). Pruébese que la matriz de φ_z en la base canónica $(1, 0)$, $(0, 1)$ de $\mathbb{R}^2$ está dada por

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix}.$$

b. Pruébese que $\varphi_{z_1 z_2} = \varphi_{z_1} \circ \varphi_{z_2}$.

41. Pruébese que $\text{Re}(iz) = -\text{Im}(z)$ y que $\text{Im}(iz) = \text{Re}(z)$ para todos los números complejos z.

42. Si $z = x + iy$, pruébese que $|x| + |y| \leq \sqrt{2}|z|$.

43. Si $a, b \subset \mathbb{C}$, demuéstrese la ***identidad del paralelogramo***:

$$|a - b|^2 + |a + b|^2 = 2(|a|^2 + |b|^2).$$

44. Demuéstrese la identidad de Lagrange para números complejos:

$$\left|\sum_{k=1}^{n} z_k w_k\right|^2 = \left(\sum_{k=1}^{n} |z_k|^2\right)\left(\sum_{k=1}^{n} |w_k|^2\right) - \left(\sum_{k<j} |z_k\overline{w}_j - z_j\overline{w}_k|^2\right).$$

De la demostración dedúzcase la desigualdad de Cauchy.

45. Pruébese que si $|z| > 1$ entonces $\lim_{n\to\infty} z^n/n = \infty$.

46. Demuéstrese que cualquier conjunto S no vacío acotado superiormente tiene un supremo, que se construye de la siguiente manera: se eligen $x_0 \in S$ y M_0 una cota superior y se define $a_0 = (x_0 + M_0)/2$; si a_0 es una cota superior, sean $M_1 = a_0$ y $x_1 = x_0$; de lo contrario, sean $M_1 = M_0$ y $x_1 > a_0$ tal que $x_1 \in S$, y se repite el procedimiento para generar las sucesiones x_n y M_n. Pruébese que ambas convergen al $\sup(S)$.

47. ***La recta larga*** Construyamos un cuerpo ordenado no arquimediano de la siguiente manera: sea $\mathbb{F}$ la unión de dos copias distintas de $\mathbb{R}$; para distinguirlas, a la segunda le ponemos primas. Definamos $x + y$ de la manera usual, $x + y' = (x + y)'$ y $x' + y' = x + y$. Definamos xy de la manera usual, $xy' = (xy)'$ y $x'y' = xy$. Definamos $x \leq y'$ para todo par x,y y $x' \leq y'$ si $x \leq y$. Pruébese que $\mathbb{F}$ cumple los requisitos. Demuéstrese que $\mathbb{F}$ *no* es completo.

Capítulo 2
La topología del espacio euclídeo

En este capítulo comenzamos nuestro estudio de aquellas propiedades básicas de $\mathbb{R}^n$ que son importantes para el concepto de función continua. Estudiaremos los conjuntos abiertos, que generalizan los intervalos abiertos de $\mathbb{R}$, así como los conjuntos cerrados, que generalizan los intervalos cerrados. El estudio de conjuntos abiertos y cerrados constituye el principio de la topología. Este estudio continuará en el capítulo 3.

La mayor parte del material del presente capítulo se basa únicamente en las propiedades fundamentales de la función distancia y, por lo tanto, tiene sentido en un espacio métrico general. Recordemos que la función distancia d para $\mathbb{R}^n$ está dada por

$$d(x,y) = \left\{ \sum_{i=1}^{n} (x_i - y_i)^2 \right\}^{1/2}$$

y que las propiedades básicas de d son

1. $d(x,y) \geq 0$.

2. $d(x,y) = 0$ sii $x = y$.

3. $d(x,y) = d(y,x)$.

4. $d(x,y) \leq d(x,z) + d(z,y)$ (desigualdad triangular).

Recordemos también que un conjunto M con una función distancia d que satisfaga estas propiedades recibe el nombre de espacio métrico.

§2.1 Conjuntos abiertos

Con el fin de definir los conjuntos abiertos, introducimos en primer lugar el concepto de disco de radio ε en un espacio métrico. Como de costumbre, nuestro ejemplo principal será $\mathbb{R}^n$.

2.1.1 Definición *Sea (M, d) un espacio métrico. Para cada $x \in M$ y $\varepsilon > 0$, el conjunto*

$$D(x, \varepsilon) = \{y \in M \mid d(x, y) < \varepsilon\}$$

*se denomina **disco de radio** ε centrado en x (o también **vecindad de radio** ε o **bola de radio** ε centrada en x). (Véase la figura 2.1-1.) Un conjunto $A \subset M$ es **abierto** si para cada $x \in A$ existe $\varepsilon > 0$ tal que $D(x, \varepsilon) \subset A$. Una **vecindad** de un punto en M es un conjunto abierto que contiene dicho punto.*

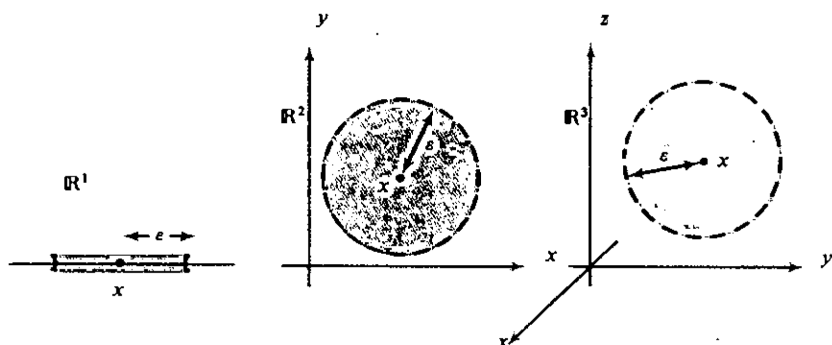


FIGURA 2.1-1 El disco de radio ε

Obsérvese que el conjunto vacío $\varnothing$ y el conjunto total M son abiertos.

Es importante observar que el valor ε de la definición de un conjunto abierto puede depender de x. Por ejemplo, el cuadrado unidad en $\mathbb{R}^2$ sin incluir su "frontera" es abierto, pero las vecindades de radio ε son cada vez más pequeñas al aproximarnos a la frontera; sin embargo, ε no puede anularse para ningún x. Véase la figura 2.1-2.

Consideremos un intervalo abierto en $\mathbb{R} = \mathbb{R}^1$, como el intervalo $]0, 1[$. Éste es, de hecho, un conjunto abierto (véase la figura 2.1-3); sin embargo, si vemos este conjunto dentro de $\mathbb{R}^2$ (como un subconjunto del eje x), ya no es abierto. Así, para que un conjunto sea abierto, es esencial especificar el espacio $\mathbb{R}^n$ o, en general, el espacio métrico del que va a considerarse como subconjunto.

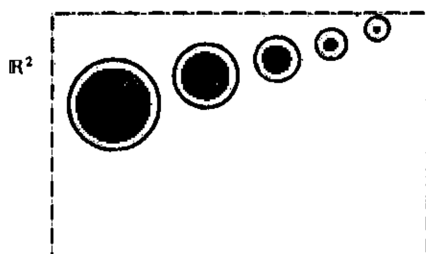


FIGURA 2.1-2 Un conjunto abierto

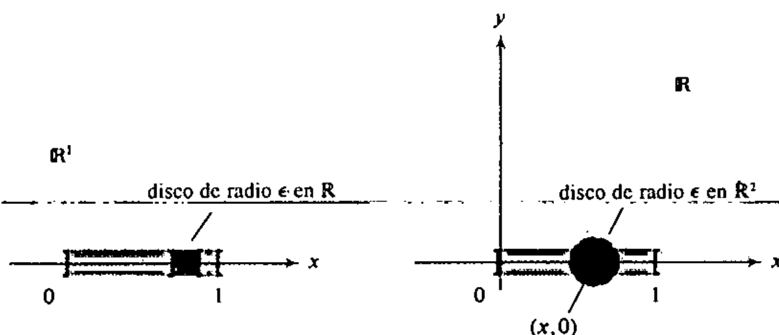


FIGURA 2.1-3]0, 1[es abierto en $\mathbb{R}$ pero no en $\mathbb{R}^2$

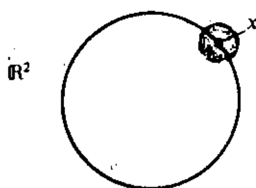


FIGURA 2.1-4 Un conjunto no abierto

Hay muchos ejemplos de conjuntos que no son abiertos. El disco unidad $\{x \in \mathbb{R}^2 \mid \|x\| \leq 1\}$ es uno de estos ejemplos. Este conjunto no es abierto, ya que si un punto está en la "frontera" (es decir, un punto x tal que $\|x\| = 1$), todo disco de radio ε contiene puntos que no están en el conjunto. (Véase la figura 2.1-4.)

2.1.2 Proposición *En un espacio métrico, todo disco de radio ε, $D(x, \varepsilon)$, es abierto.*

La idea principal de la demostración se muestra en la figura 2.1-5, en la cual se puede ver que el tamaño del disco con centro en el punto $y \in D(x, \varepsilon)$ es cada vez menor a medida que y se acerca a la frontera. Esta imagen debe clarificar "intuitivamente" el teorema.

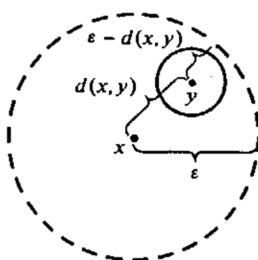


FIGURA 2.1-5 Algunas ideas útiles para la demostración de la proposición **2.1.2**

A continuación damos algunas propiedades básicas de los conjuntos abiertos:

2.1.3 Proposición *En un espacio métrico (M, d),*

i. *La intersección de un número finito de subconjuntos abiertos de M es abierta.*

ii. *La unión de una colección arbitraria de subconjuntos abiertos de M es abierta.*

iii. *El conjunto vacío $\varnothing$ y el espacio total M son abiertos.*

Para apreciar la diferencia entre las afirmaciones **i** y **ii**, obsérvese que la intersección de una familia arbitraria de conjuntos abiertos no tiene que ser abierta. Por ejemplo, en $\mathbb{R}^1$, un punto (que no es un conjunto abierto) es la intersección de todos los conjuntos abiertos que lo contienen (¿por qué?)

Nota. A un conjunto con una colección dada de subconjuntos (llamados, por definición, conjuntos abiertos) que cumplan las condiciones de la proposición **2.1.3**, y que contenga al conjunto vacío y al espacio total, se le denomina ***espacio topológico***. En este libro no estudiaremos los espacios topológicos generales, sino que trataremos principalmente los casos de $\mathbb{R}^n$ y de los espacios métricos. Sin embargo, gran parte de lo dicho en los capítulos 2, 3 y 4 se aplica también a este planteamiento más general.

2.1.4 Ejemplo *Sea $S = \{(x,y) \in \mathbb{R}^2 \mid 0 < x < 1\}$. Muéstrese que S es abierto.*

Solución En la figura 2.1-6 vemos que podemos trazar el disco de radio $r = \text{mín}\{x, 1 - x\}$ centrado en cada punto $(x, y) \in S$ y que éste está contenido completamente en S; puede verse esto en la figura y demostrarse mediante la desigualdad triangular. Por la definición, esto implica que S es abierto. ♦

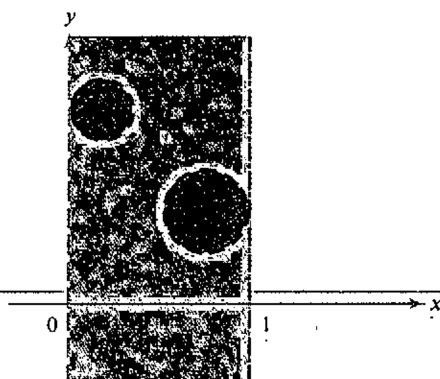


FIGURA 2.1-6 Cualquier punto de este conjunto tiene un disco en torno de él que también está dentro del conjunto

2.1.5 Ejemplo *Sea $S = \{(x, y) \in \mathbb{R}^2 \mid 0 < x \leq 1\}$. ¿Es S un conjunto abierto?*

Solución No, pues todo disco con centro en $(1, 0) \in S$ contiene puntos $(x, 0)$ tales que $x > 1$. ♦

2.1.6 Ejemplo *Sean $A \subset \mathbb{R}^n$ un conjunto abierto y $B \subset \mathbb{R}^n$. Definamos*

$$A + B = \{x + y \in \mathbb{R}^n \mid x \in A \quad y \quad y \in B\}.$$

Demuéstrese que A + B es abierto.

Solución Sea $w \in A + B$; entonces, existen puntos $x \in A$ e $y \in B$ tales que $w = x + y$. Como A es abierto, existe $\varepsilon > 0$ tal que $D(x, \varepsilon) \subset A$. Afirmamos que $D(w, \varepsilon) \subset A + B$. Supongamos que $z \in D(w, \varepsilon)$. Entonces $\|z - w\| = \|z - (x + y)\| < \varepsilon$. Pero $\|z - (x + y)\| =$

$\|(z-y)-x\|$ por lo que $z-y \in D(x, \varepsilon) \subset A$. Como $y \in B$, esto obliga a $z=(z-y)+y$ a estar en $A+B$. Así, $D(w, \varepsilon) \subset A+B$ y entonces $A+B$ es un conjunto abierto. ■

Obsérvese que hemos demostrado en realidad que si A es abierto, entonces $A+\{y\}$ es abierto. Esto implica que $A+B=\cup_{y \in B}(A+\{y\})$ también es abierto.

2.1.7 Ejemplo *Sea M un conjunto arbitrario y definamos la **métrica discreta** d_0 en M mediante $d_0(x, y)=0$ si $x=y$ y $d_0(x, y)=1$ si $x \neq y$. Demuéstrese que todo conjunto $A \subset M$ es abierto.*

Solución Dado $x \in A$, $D(x, 1/2)=\{x\} \subset A$; por definición, esto implica que A es abierto. ♦

Ejercicios de §2.1

1. Muéstrese que $\mathbb{R}^2 \backslash \{(0,0)\}$ es abierto en $\mathbb{R}^2$.
2. Sea $S=\{(x, y) \in \mathbb{R}^2 \mid xy>1\}$. Muéstrese que S es abierto.
3. Sean $A \subset \mathbb{R}$ un conjunto abierto y $B \subset \mathbb{R}^2$ definido como $B=\{(x, y) \in \mathbb{R}^2 \mid x \in A\}$. Muéstrese que B es abierto.
4. Sea $B \subset \mathbb{R}^n$ un conjunto arbitrario. Definamos $C=\{x \in \mathbb{R}^n \mid d(x, y)<1$ para algún $y \in B\}$. Muéstrese que C es abierto.
5. Sean $A \subset \mathbb{R}$ un conjunto abierto y $B \subset \mathbb{R}$. Definamos $AB=\{xy \in \mathbb{R} \mid x \in A$ e $y \in B\}$. ¿Es AB necesariamente abierto?
6. Muéstrese que $\mathbb{R}^2$ con la métrica del taxi tiene los mismos conjuntos abiertos que $\mathbb{R}^2$ con la métrica usual.

§2.2 Interior de un conjunto

2.2.1 Definición *Sean M un espacio métrico y $A \subset M$. Un punto $x \in A$ es un **punto interior** de A si existe un conjunto abierto U tal que $x \in U \subset A$. El **interior de** A es la colección de todos los puntos interiores de A y se denota* int(A)*. Este conjunto puede ser vacío.*

La condición sobre x es equivalente a la siguiente: existe $\varepsilon > 0$ tal que $D(x, \varepsilon) \subset A$ (véase el ejercicio 5 al final del capítulo). Por ejemplo, el interior de un único punto en $\mathbb{R}^n$ es vacío. El interior del disco unidad en $\mathbb{R}^2$, con su frontera, es el disco unidad sin su frontera.

Podemos dar una descripción un poco diferente del interior de un conjunto como sigue. El interior de A es la unión de todos los subconjuntos abiertos de A (pediremos al lector que lo demuestre en el ejercicio 23 al final de este capítulo). Así, tanto por la proposición **2.1.3** como directamente, $\text{int}(A)$ es abierto. Por lo tanto, $\text{int}(A)$ *es el mayor subconjunto abierto de* A. En consecuencia, si A no tiene subconjuntos abiertos, entonces $\text{int}(A) = \varnothing$. Además, es evidente que A *es abierto sii* $\text{int}(A) = A$.

2.2.2 Ejemplo *Sea* $S = \{(x,y) \in \mathbb{R}^2 \mid 0 < x \leq 1\}$. *Determínese* $\text{int}(S)$.

Solución Para determinar los puntos interiores, buscamos los puntos para los cuales podemos trazar un disco de radio ε con centro en ellos y que esté totalmente contenido en S. En la figura 2.1-6, podemos ver que éstos son los puntos (x, y) tales que $0 < x < 1$. Así, $\text{int}(S) = \{(x, y) \mid 0 < x < 1\}$. ♦

2.2.3 Ejemplo *¿Es cierto que* $\text{int}(A) \cup \text{int}(B) = \text{int}(A \cup B)$?

Solución No. En la recta real, sean $A = [0, 1]$ y $B = [1, 2]$. Entonces $\text{int}(A) =]0, 1[$ (¿por qué?) e $\text{int}(B) =]1, 2[$, de modo que $\text{int}(A) \cup \text{int}(B) =]0, 1[\cup]1, 2[=]0, 2[\setminus \{1\}$, mientras que $\text{int}(A \cup B) = \text{int}[0, 2] =]0, 2[$. ♦

2.2.4 Ejemplo *¿Es cierto en un espacio métrico general* (M, d) *que* $\text{int}\{y \in M \mid d(y, x_0) \leq r\} = \{y \in M \mid d(y, x_0) < r\}$ *dados* $x_0 \in M$ *y* $r > 0$?

Solución No. Por ejemplo, consideremos un conjunto M con la métrica discreta, $x_0 \in M$ y $r = 1$. Entonces $\{y \in M \mid d(y, x_0) \leq 1\} = M$, por lo que su interior es todo M. Por otro lado, $\{y \mid d(y, x_0) < 1\} = \{x_0\}$, que no es M si éste tiene más de un punto. ♦

Ejercicios de §2.2

1. Sea $S = \{(x, y) \in \mathbb{R}^2 \mid xy \geq 1\}$. Determínese $\text{int}(S)$.

2. Sea $S = \{(x, y, z) \in \mathbb{R}^3 \mid 0 < x < 1, y^2 + z^2 \leq 1\}$. Determínese $\text{int}(S)$.

3. Si $A \subset B$, ¿es cierto que $\text{int}(A) \subset \text{int}(B)$?

4. ¿Es cierto que $\text{int}(A) \cap \text{int}(B) = \text{int}(A \cap B)$?

5. Sean (M, d) un espacio métrico, $x_0 \in M$ y $r > 0$. Muéstrese que

$$D(x_0, r) \subset \text{int}\{y \in M \mid d(y, x_0) \leq r\}.$$

§2.3 Conjuntos cerrados

2.3.1 Definición *Un conjunto B en un espacio métrico M es **cerrado** si su complementario (es decir, el conjunto $M \backslash B$) es abierto.*

Por ejemplo, un punto en $\mathbb{R}^n$ es un conjunto cerrado. El conjunto en $\mathbb{R}^2$ formado por el disco unidad y su círculo frontera es cerrado. En pocas palabras, podemos decir que un conjunto es cerrado si contiene a sus "puntos frontera" (precisaremos este concepto intuitivo en §**2.6**). Véase la figura 2.3-1. La intuición es difícil de utilizar para algunos conjuntos complicados, por lo que en esos casos se echa mano de la definición técnica **2.3.1**.



FIGURA 2.3-1 Ejemplos de conjuntos cerrados

Es posible tener un conjunto que no sea abierto ni cerrado. Por ejemplo, en $\mathbb{R}^1$, un intervalo semiabierto $]0, 1]$ no es abierto ni cerrado. Así, si sabemos que un conjunto A no es abierto, *no podemos* concluir que sea necesariamente cerrado.

El siguiente teorema es análogo a la proposición **2.1.3**.

2.3.2 Proposición *En un espacio métrico (M, d),*

i. *La unión de un número finito de subconjuntos cerrados es cerrada.*

ii. *La intersección de una familia arbitraria de subconjuntos cerrados es cerrada.*

iii. *El espacio total M y el conjunto vacío ∅ son cerrados.*

Este teorema se sigue de la proposición **2.1.3**, teniendo en cuenta que uniones e intersecciones se intercambian cuando tomamos complementarios (véase la Introducción). Dejamos la demostración al lector (ejercicio 22 al final del capítulo), quien también debe probar que no es posible reemplazar en la primera afirmación la unión finita por una unión arbitraria.

2.3.3 Ejemplo *Sea* $S = \{(x, y) \in \mathbb{R}^2 \mid 0 < x \leq 1, 0 \leq y \leq 1\}$. *¿Es S cerrado?*

Solución Véase la figura 2.3-2. Intuitivamente, S no es cerrado, pues la parte de su frontera que está en el eje x no está en S. Además, su complementario no es abierto, pues cualquier disco de radio ε con centro en un punto del eje y, como el punto $(0, 1/2)$, intersecará a S (y por lo tanto no estará contenido en $\mathbb{R}^2 \setminus S$). ♦

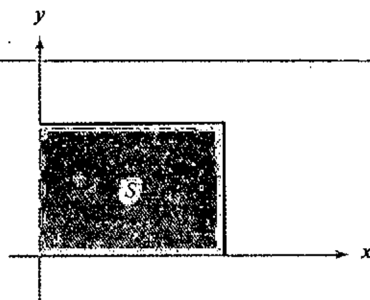


FIGURA 2.3-2 ¿Es cerrado este conjunto?

2.3.4 Ejemplo *Sea* $S = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$. *¿Es S cerrado?*

Solución Sí, S es el disco unidad, con su frontera. El complementario es un conjunto abierto, pues para $(x, y) \in \mathbb{R}^2 \setminus S$, el disco de radio $\varepsilon = \sqrt{x^2 + y^2} - 1$ está totalmente contenido en $\mathbb{R}^2 \setminus S$. (Figura 2.3-3). ♦

2.3.5 Ejemplo *Muéstrese que cualquier conjunto finito en* $\mathbb{R}^n$ *es cerrado.*

Solución Los puntos (individuales) son cerrados, por lo que la afirmación es una consecuencia de la proposición **2.3.2i**.

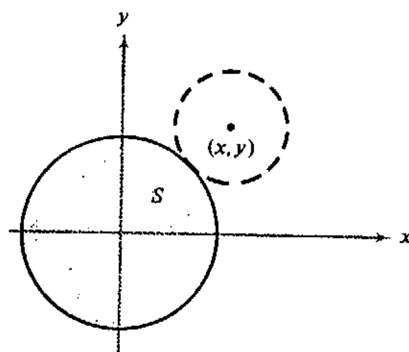


FIGURA 2.3-3 El complementario del conjunto S en el ejemplo **2.3.4** es abierto

2.3.6 Ejemplo *Sean (M, d) un espacio métrico y $A \subset M$ un conjunto finito. Sea* $B = \{x \in M \mid d(x, y) \leq 1$ para algún $y \in A\}$. *Muéstrese que B es cerrado.*

Solución Mostraremos que $M \backslash B$ es abierto. Sea $z \in M \backslash B$, de modo que $d(z, y) > 1$ para todo $y \in A$. Sean $y_1, \ldots, y_N$ los puntos de A; entonces, $d(z, y_i) > 1$ para $i = 1, \ldots, N$. Sea ε el mínimo de los números $d(z, y_i) - 1, \ldots, d(z, y_N) - 1$, por lo que $\varepsilon > 0$. Si $d(x, z) < \varepsilon/2$, la desigualdad triangular muestra que $d(x, y_i) + d(x, z) \geq d(y_i, z)$ o $d(x, y_i) \geq d(y_i, z) - \varepsilon/2$, lo cual, por la construcción de ε, es estrictamente mayor que 1. Así, $D(z, \varepsilon) \subset M \backslash B$, luego $M \backslash B$ es abierto y B es cerrado. ♦

Ejercicios de §2.3

1. Sea $S = \{(x, y) \in \mathbb{R}^2 \mid x \geq 1 \text{ e } y \geq 1\}$. ¿Es S cerrado?

2. Sea $S = \{(x, y) \in \mathbb{R}^2 \mid x = 0, 0 < y < 1\}$. ¿Es S cerrado?

3. Repítase el ejemplo **2.3.5** de forma directa, mostrando esta vez que el complementario es abierto.

4. Sea $A \subset \mathbb{R}^n$ un conjunto arbitrario. Muéstrese que $\mathbb{R}^n \backslash (\text{int}\, A)$ es cerrado.

5. Sea $S = \{x \in \mathbb{R} \mid x \text{ es irracional}\}$. ¿Es S cerrado?

6. Proporciónese una solución alternativa al ejemplo **2.3.6**, mostrando que B es la unión de conjuntos cerrados.

§2.4 Puntos de acumulación

Otra forma útil para determinar si un conjunto es cerrado o no se basa en el concepto de punto de acumulación.

2.4.1 Definición *Un punto x en un espacio métrico M es un* ***punto de acumulación*** *de un conjunto $A \subset M$ si todo conjunto abierto U que contiene a x contiene también algún punto de A, distinto de x.*

En otras palabras, un punto de acumulación de un conjunto A es un punto para el cual hay puntos de A arbitrariamente cercanos a él. A los puntos de acumulación se les denomina a veces ***puntos límite***. La proposición **2.1.2** nos permite afirmar que x es un punto de acumulación de A si y sólo si *para cada $\varepsilon > 0$, $D(x, \varepsilon)$ contiene algún punto y de A tal que $y \neq x$*. Por ejemplo, en $\mathbb{R}^1$, un conjunto formado por un único punto no tiene puntos de acumulación, mientras que para el intervalo abierto $]0, 1[$ todos los puntos de $[0, 1]$ son puntos de acumulación. Obsérvese que un punto de acumulación de un conjunto *no tiene por qué* pertenecer a él. Las definiciones de punto de acumulación y de conjunto cerrado tienen una estrecha relación, como lo muestra el siguiente teorema:

2.4.2 Teorema *Un conjunto $A \subset M$ es cerrado sii todos los puntos de acumulación de A pertenecen a A.*

Un conjunto no tiene por qué tener puntos de acumulación (un único punto y el conjunto de enteros en $\mathbb{R}^1$ son ejemplos de ello); pero entonces el teorema **2.4.2** se aplica trivialmente y de este modo podemos concluir que un conjunto de este tipo es cerrado.

El teorema **2.4.2** es intuitivamente claro, pues la propiedad de ser cerrado significa, *grosso modo*, que un conjunto contiene todos los puntos de su "frontera", y tales puntos son puntos de acumulación. Este tipo de argumento es impreciso y tiene ciertas dificultades, por lo que debemos tener cuidado: algunos conjuntos son lo bastante complejos como para que falle nuestra intuición. Por ejemplo, considérese $A = \{1/n \in \mathbb{R} \mid n = 1, 2, 3, \ldots\} \cup \{0\}$. Éste es un conjunto cerrado (¡compruébese!) y su único punto de acumulación es $\{0\}$, que está en A. Pero nuestra intuición de "frontera" no es clara para este conjunto; por lo tanto, necesitamos argumentos más cuidadosos.

2.4.3 Ejemplo *Sea $S = \{x \in \mathbb{R} \mid x \in [0, 1]$ y x es racional$\}$. Encuéntrense los puntos de acumulación de S.*

Solución El conjunto de puntos de acumulación está formado por todos los puntos de $[0, 1]$. De hecho, sea $y \in [0, 1]$ y sea $D(y, \varepsilon) =]y - \varepsilon, y + \varepsilon[$ una vecindad de y.

Podemos encontrar puntos racionales en [0, 1] arbitrariamente cercanos a y (distintos de y), y en particular en $D(y, \varepsilon)$. Por lo tanto, y es un punto de acumulación. Ningún punto $y \notin [0, 1]$ es un punto de acumulación, pues existe un disco de radio ε centrado en y que lo contiene pero que no interseca a [0, 1]. ♦

2.4.4 Ejemplo *Verifíquese el teorema* **2.4.2** *para el conjunto* $A = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1 \text{ o } x = 2\}$.

Solución El conjunto A aparece en la figura 2.4-1. Es evidente tanto que A es cerrado, como que el conjunto de puntos de acumulación de A es igual a A, de modo que cada punto de acumulación de A es un elemento de A. Obsérvese que en $\mathbb{R}$, el conjunto de puntos de acumulación de $[0, 1] \cup \{2\}$ es [0, 1] *sin* el punto {2}. ♦

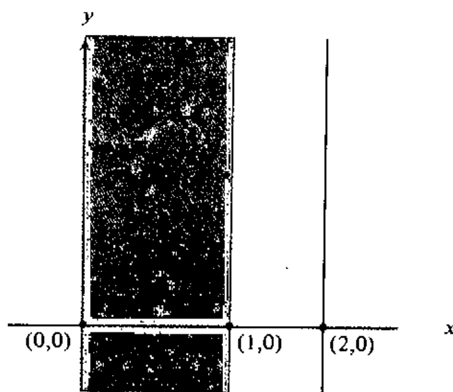


FIGURA 2.4-1 Este conjunto contiene todos sus puntos de acumulación

2.4.5 Ejemplo *Sea* $S = \{(x, y) \in \mathbb{R}^2 \mid y < x^2 + 1\}$. *Determínense los puntos de acumulación de S.*

Solución El conjunto S aparece en la figura 2.4-2. Los puntos de acumulación forman el conjunto $\{(x, y) \mid y \leq x^2 + 1\}$, como lo ilustra claramente la figura.

Ejemplo 2.4.6 *En un espacio métrico general M, sea* $B(x, r) = \{y \in M \mid d(x, y) \leq r\}$. *¿Es cierto que los puntos de B(x, r) son puntos de acumulación de D(x, r)?*

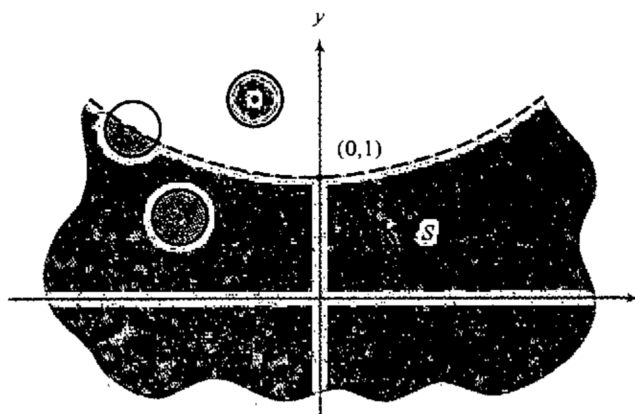


FIGURA 2.4-2 El conjunto de puntos (x, y) que satisfacen $y < x^2 + 1$

Solución Esto no es cierto en general, aunque sí lo es en $\mathbb{R}^n$. Por ejemplo, si M es un conjunto con la métrica discreta, entonces $D(x, 1) = \{x\}$, mientras que $B(x, 1) = M$. El conjunto $\{x\}$ no tiene puntos de acumulación, pues está formado por un único punto, de modo que si M tiene más de un punto, $B(x, 1)$ contendrá puntos que no son puntos de acumulación de $D(x, 1)$. ♦

2.4.7 Ejemplo *Muéstrese que una sucesión acotada x_n de puntos distintos en $\mathbb{R}$ tiene un punto de acumulación; es decir, el conjunto $\{x_1, x_2, \ldots\}$ tiene un punto de acumulación.*

Solución En el capítulo 1 (véase **1.4.3**) demostramos que existe una subsucesión (que denotamos x_{n_k}) de x_n que converge: $x_{n_k} \to x$. Entonces, x es un punto de acumulación, puesto que para cada $\varepsilon > 0$, $D(x, \varepsilon)$ contiene a x_{n_k} con n_k suficientemente grande (por definición de convergencia) y al menos uno de los términos x_{n_k} es distinto de x, ya que los x_{n_k} son todos distintos. ♦

Ejercicios de §2.4

1. Hállense los puntos de acumulación de $A = \{(x, y) \in \mathbb{R}^2 \mid y = 0 \text{ y } 0 < x < 1\}$.

2. Si $A \subset B$ y x es un punto de acumulación de A, ¿es x también un punto de acumulación de B?

3. Encuéntrense los puntos de acumulación de los siguientes conjuntos en $\mathbb{R}^2$:

a. $\{(m, n) \mid m, n \text{ enteros}\}$

b. $\{(p, q) \mid p, q \text{ racionales}\}$

c. $\{(m/n, 1/n) \mid m, n \text{ enteros}, n \neq 0\}$

d. $\{(1/n + 1/m, 0) \mid n, m \text{ enteros}, n \neq 0, m \neq 0\}$

4. Sea $A \subset \mathbb{R}$ un conjunto no vacío, acotado superiormente y $x = \sup(A)$. ¿Debe ser x un punto de acumulación de A?

5. Verifíquese el teorema **2.4.2** para el conjunto $A = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y + 2x = 3\}$.

6. Sea M un conjunto con la métrica discreta y $A \subset M$ cualquier subconjunto. Determínese el conjunto de puntos de acumulación de A.

§2.5 Clausura de un conjunto

El interior de un conjunto A es el mayor subconjunto abierto de A. De forma similar, podemos construir el menor conjunto cerrado que contiene a un conjunto A, al que llamamos *clausura de* A y denotamos $\text{cl}(A)$ o, a veces, $\overline{A}$.

2.5.1 Definición *Sean (M,d) un espacio métrico y $A \subset M$. La **clausura** de A, denotada* $\text{cl}(A)$, *es la intersección de todos los conjuntos cerrados que contienen a A.*

Puesto que la intersección de una familia *arbitraria* de conjuntos cerrados es cerrada, $\text{cl}(A)$ *es cerrado*; también es inmediato que $\text{cl}(A) \supset A$ y que *A es cerrado sii* $\text{cl}(A) = A$. Por ejemplo, en $\mathbb{R}^1$, $\text{cl}(]0, 1[) = [0, 1]$. La relación entre la clausura y los puntos de acumulación es la siguiente:

2.5.2 Proposición *Si $A \subset M$,* $\text{cl}(A)$ *es el conjunto formado por A junto con los puntos de acumulación de A; es decir,* $\text{cl}(A) = A \cup \{\text{puntos de acumulación de } A\}$.

En otras palabras, para determinar la clausura de un conjunto A, añadimos a A todos los puntos de acumulación que no se encuentren ya en A. La proposición **2.5.2** debería ser clara desde un punto de vista intuitivo, gracias a los ejemplos ya presentados.

2.5.3 Ejemplo *Determínese la clausura de* $A = [0, 1[\cup \{2\}$ *en* $\mathbb{R}$.

Solución Los puntos de acumulación son [0, 1], por lo que la clausura es $[0, 1] \cup \{2\}$, que claramente es el menor conjunto cerrado que contiene a A que podríamos encontrar. ♦

2.5.4 Ejemplo *Para cualquier* $A \subset \mathbb{R}^n$, *muéstrese que* $\mathbb{R}^n \backslash \mathrm{cl}(A)$ *es abierto.*

Solución $\mathrm{cl}(A)$ es un conjunto cerrado; por la definición de un conjunto de este tipo, su complementario es abierto. ♦

2.5.5 Ejemplo *¿Es cierto que* $\mathrm{cl}(A \cap B) = \mathrm{cl}(A) \cap \mathrm{cl}(B)$*?*

Solución No. Tomemos, por ejemplo, $A = [0, 1]$ y $B =]1, 2]$. Entonces $A \cap B = \varnothing$ y $\mathrm{cl}(A) \cap \mathrm{cl}(B) = \{1\}$. ♦

2.5.6 Ejemplo *Dado un subconjunto A de un espacio métrico, muéstrese que* $x \in \mathrm{cl}(A)$ *si y sólo si* $\inf\{d(x, y) \mid y \in A\} = 0$.

Solución En primer lugar, sean $x \in \mathrm{cl}(A)$ y $\alpha = \inf\{d(x, y) \mid y \in A\}$. Si $x \in A$, entonces tomamos $x = y$ y obtenemos $\alpha = 0$. Si x es un punto de acumulación de A, entonces para cada $\varepsilon > 0$ existe $y \in A$ tal que $d(x, y) < \varepsilon$. Así, de nuevo, $\alpha = 0$. Recíprocamente, si $\alpha = 0$ y $x \notin$ A, entonces para cada $\varepsilon > 0$ existe $y \in A$ tal que $d(x, y) < \varepsilon$ (por una propiedad del ínfimo), de modo que x es un punto de acumulación de A y entonces $x \in \mathrm{cl}(A)$. ♦

Ejercicios de §2.5

1. Determínese la clausura de $A = \{(x, y) \in \mathbb{R}^2 \mid x > y^2\}$.

2. Determínese la clausura de $\{1/n \mid n = 1, 2, 3, \ldots\}$ en $\mathbb{R}$.

3. Sea $A = \{(x, y) \in \mathbb{R}^2 \mid x \text{ es racional}\}$. Determínese $\mathrm{cl}(A)$.

4. **a.** Sea $A \subset \mathbb{R}^n$; muéstrese que $\mathrm{cl}(A) \backslash A$ está formado en su totalidad por puntos de acumulación de A.

b. ¿Tienen que ser necesariamente todos los puntos de acumulación de A?

5. En un espacio métrico general M, sea $A \subset D(x, r)$ para $x \in M$ y $r > 0$. Muéstrese que $\mathrm{cl}(A) \subset B(x, r) = \{y \in M \mid d(x, y) \leq r\}$.

§2.6 Frontera de un conjunto

Si consideramos el disco unidad en $\mathbb{R}^2$, sabemos lo que queremos denominar la frontera: la elección obvia es la circunferencia unidad. Para conjuntos más complejos, como los racionales, no está claro qué debería ser la frontera. Por lo tanto, necesitamos una definición precisa.

2.6.1 Definición *Para un conjunto dado A en un espacio métrico (M, d), la **frontera** se define como el conjunto*

$$\partial(A) = \text{cl}(A) \cap \text{cl}(M \setminus A).[1]$$

Puesto que la intersección de dos conjuntos cerrados es cerrada, ∂A es un conjunto cerrado. Obsérvese también que $\partial A = \partial(M \setminus A)$. La proposición **2.5.2** implica que la frontera tiene también la siguiente descripción.

2.6.2 Proposición *Sea $A \subset M$. Entonces $x \in \partial A$ sii para todo $\varepsilon > 0$, $D(x, \varepsilon)$ contiene puntos de A y de $M \setminus A$. (Estos puntos pueden incluir al propio x.) Véase la figura 2.6-1.*

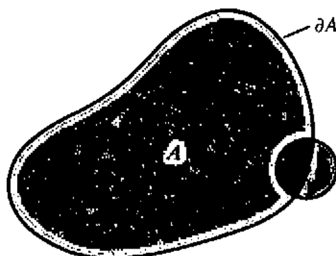


FIGURA 2.6-1 Los discos en torno a los puntos frontera contienen puntos del conjunto y puntos que no están en el conjunto

La definición original establece que ∂A es la "línea divisoria" entre A y $M \setminus A$. Esto es también lo que afirma la proposición **2.6.2**, por lo que debe ser intuitivamente clara.

[1] En el original se usa la notación bd(A), por *boundary*, frontera, en inglés. (*N. del T.*)

2.6.3 Ejemplo *Sea* $A = \{x \in \mathbb{R} \mid x \in [0, 1]$ y *x es racional*$\}$. *Determínese* ∂A.

Solución $\partial A = [0, 1]$, pues para todo $\varepsilon > 0$ y $x \in [0, 1]$, $D(x, \varepsilon) =]x - \varepsilon, x + \varepsilon[$ contiene puntos racionales e irracionales. El lector debería verificar también que $\partial A = [0, 1]$ mediante la definición original de ∂A. Este ejemplo muestra que si $A \subset B$, esto no necesariamente implica que $\partial A \subset \partial B$. Tómese A como en este ejemplo y $B = [0, 1]$ en $\mathbb{R}$.

2.6.4 Ejemplo *Si* $x \in \partial A$, *¿debe x ser necesariamente un punto de acumulación?*

Solución No. Por ejemplo, sea $A = \{0\} \subset \mathbb{R}$. Entonces A no tiene puntos de acumulación, pero $\partial A = \{0\}$. ♦

2.6.5 Ejemplo *Sea* $S = \{(x, y) \in \mathbb{R}^2 \mid x^2 - y^2 > 1\}$. *Determínese* ∂S.

Solución El conjunto S está representado en la figura 2.6-2. Claramente ∂S está formada por la hipérbola $x^2 - y^2 = 1$. ♦

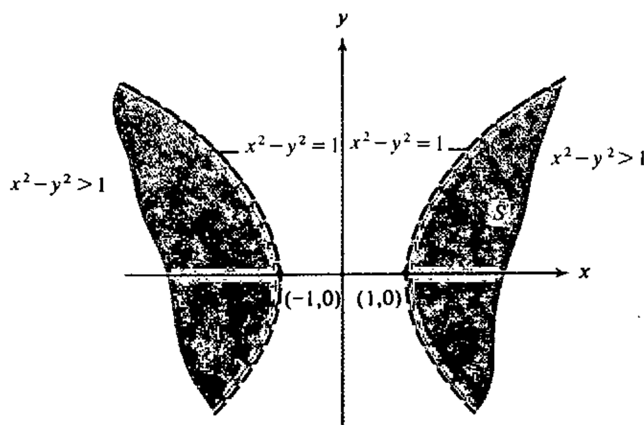


FIGURA 2.6-2 El conjunto S del ejemplo **2.6.5**

Ejercicios de §2.6

1. Determínese ∂A para el conjunto $A = \{1/n \in \mathbb{R} \mid n = 1, 2, 3, \ldots\}$.

2. Si $x \in \mathrm{cl}(A)\backslash A$, muéstrese que entonces $x \in \partial A$. ¿Es cierto el recíproco?

3. Determínese ∂A si $A = \{(x, y) \in \mathbb{R}^2 \mid x \leq y\}$.

4. ¿Es cierto siempre que $\partial A = \partial \mathrm{int}\, A$?

5. Sean $A \subset \mathbb{R}$ un conjunto acotado, no vacío y $x = \sup(A)$. ¿Está x en ∂A?

6. Demuéstrese que la frontera de un conjunto en $\mathbb{R}^2$ con la métrica usual es igual a la que se obtendría con la métrica del taxi.

§2.7 Sucesiones

La definición de convergencia de una sucesión en $\mathbb{R}^n$ es similar a la de convergencia de una sucesión de números reales. De hecho, la definición tiene perfecto sentido en un espacio métrico.

2.7.1 Definición *Sea (M, d) un espacio métrico y x_k una sucesión de puntos en M. Decimos que x_k **converge** a un punto $x \in M$ y escribimos*

$$\lim_{k\to\infty} x_k = x \quad o \quad x_k \to x \text{ cuando } k \to \infty,$$

si para todo conjunto abierto U que contiene a x existe un entero N tal que $x_k \in U$ si $k \geq N$. Véase la figura 2.7-1.

Esta definición coincide con la definición usual ε-N, como lo muestra el siguiente teorema.

2.7.2 Proposición *Una sucesión x_k en M converge a $x \in M$ sii para todo $\varepsilon > 0$ existe N tal que $k \geq N$ implica $d(x, x_k) < \varepsilon$.*

Particularicemos nuestro estudio de la convergencia al caso de $\mathbb{R}^n$, donde la definición **2.7.1** se lee: una sucesión de puntos v_k en $\mathbb{R}^n$ converge a v si para todo $\varepsilon > 0$ existe N tal que $\|v_k - v\| < \varepsilon$ si $k \geq N$. De hecho, esto se debe a que $d(v, v_k) = \|v - v_k\|$. Obsérvese además que para $n = 1$ recuperamos la definición de convergencia en $\mathbb{R}$. La definición de convergencia en cualquier otro espacio normado es esencialmente la misma: si *V es*

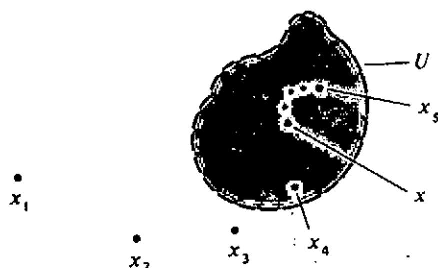


FIGURA 2.7-1 Convergencia de una sucesión en $\mathbb{R}^2$

un espacio normado y v_k es una sucesión en V, entonces v_k ***converge*** *a $v \in V$ si $\|v_k - v\| \to 0$ cuando $k \to \infty$.* Para $\mathbb{R}^2$, la situación se ilustra en la figura 2.7-1. Gran parte de lo que ya hemos hecho en el caso de dimensión 1 sigue siendo válido en $\mathbb{R}^n$, pero hay otra parte que no. Lo más destacable es que no existe un orden natural impuesto sobre $\mathbb{R}^n$, por lo que no podemos aplicar el estudio de las sucesiones monótonas a este caso general. La mayor parte del resto se obtiene reemplazando los valores absolutos por normas o distancias en todas partes. Por ejemplo, puesto que los espacios normados son espacios vectoriales, sabemos sumar vectores y multiplicarlos por números, por lo que podemos analizar la aritmética de las sucesiones.

2.7.3 Proposición *Sean v_k y w_k sucesiones de vectores en un espacio normado (como por ejemplo $\mathbb{R}^n$). λ_k una sucesión de números en $\mathbb{R}$, $\lambda \in \mathbb{R}$ una constante y u un vector constante. Si $v_k \to v$, $w_k \to w$ y $\lambda_k \to \lambda$, entonces*

i. $v_k + w_k \to v + w$.

ii. $\lambda v_k \to \lambda v$.

iii. $\lambda_k u \to \lambda u$.

iv. $\lambda_k v_k \to \lambda v$.

v. *Si $\lambda_k \neq 0$ y $\lambda \neq 0$, entonces $\frac{1}{\lambda_k} v_k \to \frac{1}{\lambda} v$.*

Para algunos cálculos emplearemos la representación específica de los vectores en $\mathbb{R}^n$ en términos de un número finito de coordenadas. Si v y v_k están en $\mathbb{R}^n$, entonces escribimos $v = (v^1, \ldots, v^n)$ y $v_k = (v_k^1, v_k^2, \ldots, v_k^n)$, donde v^i, $v_k^j \in \mathbb{R}$.

2.7.4 Proposición *$v_k \to v$ en $\mathbb{R}^n$ si y sólo si cada sucesión de coordenadas converge a la coordenada correspondiente de v como una sucesión en $\mathbb{R}$. Es decir, $\lim_{k\to\infty} v_k = v$ en $\mathbb{R}^n$ sii $\lim_{k\to\infty} = v_k^i = v^i$ en $\mathbb{R}$ para cada $i = 1, 2, \ldots, n$.*

Esto puede escribirse de una forma más compacta como

$$\lim_{k\to\infty}(v_k^1,\ldots,v_k^n)=\left(\lim_{k\to\infty}v_k^1,\ldots,\lim_{k\to\infty}v_k^n\right).$$

Para el caso de $\mathbb{R}^2$, la proposición **2.7.4** debería ser evidente de la figura 2.7-1. Obsérvese que esta proposición no tiene sentido para un espacio métrico general.

2.7.5 Ejemplo *Muéstrese que la sucesión de vectores $v_k = (1/k, 1/k^2)$ converge a $(0, 0)$ en $\mathbb{R}^2$ cuando $k \to \infty$.*

Solución Cada una de las sucesiones componentes $1/k$ y $1/k^2$ converge a 0. Por la proposición **2.7.4**, los vectores $(1/k, 1/k^2)$ convergen a $(0, 0)$ en $\mathbb{R}^2$. ♦

Las sucesiones pueden emplearse para determinar si un conjunto es cerrado. El método es el siguiente:

2.7.6 Proposición

i. *Un conjunto $A \subset M$ es cerrado sii para cada sucesión $x_k \in A$ que converge en M, el límite es un elemento de A.*

ii. *Para un conjunto $B \subset M$, $x \in \mathrm{cl}(B)$ sii existe una sucesión $x_k \in B$ tal que $x_k \to x$.*

Debemos hacer notar que las sucesiones en **i** e **ii** pueden ser triviales; es decir, se permite que $x_k = x$ para todo k.

2.7.7 Ejemplo *Sea $x_n \in \mathbb{R}^m$ una sucesión convergente tal que $\|x_n\| \le 1$ para todo n. Muéstrese que el límite x también satisface $\|x\| \le 1$. Si $\|x_n\| < 1$, ¿debe cumplirse entonces que $\|x\| < 1$?*

Solución La bola unidad $B = \{y \in \mathbb{R}^m \mid \|y\| \le 1\}$ es cerrada. Por lo tanto, por la proposición **2.7.6i**, $x_n \in B$ y $x_n \to x$ implican que $x \in B$. Esto no es cierto si reemplazamos $\le$ por $<$; por ejemplo, considérese la sucesión $x_n = 1 - (1/n)$ en $\mathbb{R}$. ♦

2.7.8 Ejemplo *Determínese la clausura de $A = \{1/n \in \mathbb{R} \mid n = 1, 2, \ldots\}$.*

Solución Podemos usar, por ejemplo, la proposición **2.7.6ii**. La sucesión $1/n \to 0$, por lo que $0 \in \mathrm{cl}(A)$. Tomando otras sucesiones de A no se obtiene ningún punto nuevo, por lo que $\mathrm{cl}(A) = A \cup \{0\}$. ♦

Ejercicios de §2.7

1. Determínese el límite de la sucesión $((\mathrm{sen}\ n)^n/n,\ 1/n^2)$ en $\mathbb{R}^2$.

2. Sea $x_n \to x$ en $\mathbb{R}^m$. Muéstrese que $A = \{x_n \mid n = 1, 2 \ldots\} \cup \{x\}$es cerrado.

3. Sean $A \subset \mathbb{R}^m$, $x_n \in A$, y $x_n \to x$. Muéstrese que $x \in \mathrm{cl}(A)$.

4. Verifíquese la proposición **2.7.6ii** para el conjunto $B = \{(x, y) \in \mathbb{R}^2 \mid x < y\}$.

5. Sea $S = \{x \in \mathbb{R} \mid x \text{ es racional y } x^2 < 2\}$. Calcúlese $\mathrm{cl}(S)$.

§2.8 Completitud

En un espacio métrico general, o incluso en $\mathbb{R}^n$, los conceptos de sucesión monótona y supremo no tienen sentido. Sin embargo, la caracterización de la completitud por medio de las sucesiones de Cauchy sí lo tiene.

2.8.1 Definición *Sea (M, d) un espacio métrico. Una* ***sucesión de Cauchy*** *es una sucesión $x_k \in M$ tal que para todo $\varepsilon > 0$, existe N tal que si $k, l \geq N$, entonces $d(x_k, x_l) < \varepsilon$. El espacio M es* ***completo*** *sii toda sucesión de Cauchy en M converge a un punto de M.*

En un espacio normado, como $\mathbb{R}^n$, una sucesión v_k es una sucesión de Cauchy si para todo $\varepsilon > 0$ existe N tal que $\|v_k - v_j\| < \varepsilon$ siempre que $k, j \geq N$.

2.8.2 Definición *Una sucesión x_k en un espacio normado está* ***acotada*** *si existe un número C tal que $\|x_k\| \leq C$ para todo k. En un espacio métrico se requiere que exista un punto x_0 tal que $d(x_k, x_0) \leq C$ para todo k.*

2.8.3 Proposición *Una sucesión convergente en un espacio normado o métrico está acotada.*

Las siguientes son algunas propiedades generales de las sucesiones de Cauchy; compárese con lo expuesto en §**1.4.**

2.8.4 Proposición

i. *Toda sucesión convergente en un espacio métrico es una sucesión de Cauchy.*

ii. *Una sucesión de Cauchy en un espacio métrico debe estar acotada.*

iii. *Si una subsucesión de una sucesión de Cauchy converge a x, entonces la sucesión converge a x.*

Por el teorema **1.4.4**, $\mathbb{R}$ es un *espacio métrico completo.* Un ejemplo de un espacio que no es completo es el conjunto de los números racionales, con $d(x, y) = |x - y|$. Otro ejemplo es $\mathbb{R}\backslash\{0\}$ con la misma métrica. El teorema **2.8.5** afirma que $\mathbb{R}^n$ *también es un espacio métrico completo.*

2.8.5 Teorema *Una sucesión x_k en $\mathbb{R}^n$ converge a un punto en $\mathbb{R}^n$ sii es una sucesión de Cauchy.*

Este teorema es una consecuencia del resultado correspondiente para $\mathbb{R}$ y la proposición **2.7.4**, pues la convergencia en $\mathbb{R}^n$ es equivalente a la convergencia componente a componente en $\mathbb{R}$.

El concepto de punto límite nos ayudará a desarrollar el concepto de completitud. Vamos a generalizar la definición y algunas de las propiedades básicas que empleamos en §**1.5.**

2.8.6 Definición *Un punto x en un espacio métrico es un **punto límite** de la sucesión x_k si para todo $\varepsilon > 0$ hay infinitos valores de k tales que $d(x_k, x) < \varepsilon$.*

2.8.7 Proposición *Si x_k es una sucesión en un espacio métrico M y $x \in M$, entonces*

i. *x es un punto límite sii para cada $\varepsilon > 0$ y cada entero N existe $k > N$ tal que $d(x_k, x) < \varepsilon$.*

ii. *x es un punto límite sii existe una subsucesión convergente a x.*

iii. *$x_k \to x$ sii toda subsucesión converge a x.*

iv. *$x_k \to x$ sii toda subsucesión de x_k tiene a su vez una subsucesión que converge a x.*

2.8.8 Ejemplo *Sea (M, d) un espacio métrico completo y $N \subset M$ un subconjunto cerrado. Muéstrese que N también es completo.*

Solución Se entiende que la métrica utilizada en N es la misma que la de M. Sea x_n una sucesión de Cauchy en N, la cual, considerada como una sucesión en M, sigue siendo una sucesión de Cauchy. Como M es completo, esta sucesión converge en M a un punto x. Como $x_n \in N$, $x_n \to x$ y N es cerrado, $x \in N$. Así, x_n converge en N y, por tanto, N es completo. ♦

2.8.9 Ejemplo *Muéstrese que el conjunto de puntos límite de una sucesión x_k es cerrado.*

Solución Sea y_n una sucesión de puntos límite tales que $y_n \to y$. Para cada $\varepsilon > 0$, existe N tal que $d(y_n, y) < \varepsilon/2$ si $n \geq N$. Hay infinitos puntos x_k tales que $d(y_k, y_N) < \varepsilon/2$, pues y_N es un punto límite. La desigualdad triangular implica que $d(x_k, y) \leq d(x_k, y_N) + d(y_N, y) < \varepsilon/2 + \varepsilon/2 = \varepsilon$, por lo que y es también un punto límite de x_k. ♦

Ejercicios de §2.8

1. Si (M, d) es un espacio métrico y $N \subset M$ es completo, muéstrese que N es cerrado.

2. Sea (M, d) un espacio métrico con la propiedad de que toda sucesión acotada tiene una subsucesión convergente. Demuéstrese que M es completo.

3. Sea M un conjunto con la métrica discreta. ¿Es M completo?

4. Muéstrese que una sucesión de Cauchy sólo puede tener como mucho un punto límite.

5. Supongamos que un espacio métrico tiene la propiedad de que toda sucesión acotada tiene al menos un punto límite. Demuéstrese que M es completo.

§2.9 Series de números reales y de vectores

Al igual que en $\mathbb{R}$, podemos estudiar series en $\mathbb{R}^n$ o, de forma más general, en un espacio normado.

2.9.1 Definición *Sea V un espacio normado. Una serie $\sum_{k=0}^{\infty} x_k$, donde $x_k \in V$,* ***converge a*** *$x \in V$ si la sucesión de sumas parciales $s_k = \sum_{i=0}^{k} x_i$ converge a x; en ese caso, escribimos $\sum_{k=0}^{\infty} x_k = x$.*

Como en nuestro estudio de las sucesiones en §**2.7**, $\sum_{k=0}^{\infty} x_k = x$ equivale a que las series componentes converjan a las correspondientes componentes de x. Si aplicamos el criterio de Cauchy a s_k obtenemos el siguiente resultado:

2.9.2 Teorema *Sea V un espacio normado completo (como, por ejemplo, $\mathbb{R}^n$). Una serie $\sum x_k$ en V converge sii para todo $\varepsilon > 0$ existe N tal que $k \geq N$ implique*

$$\|x_k + x_{k+1} + \cdots + x_{k+p}\| < \varepsilon \text{ para todo entero } p = 0, 1, 2, \ldots.$$

En particular, si tomamos $p = 0$, vemos que si $\sum x_k$ converge, entonces $x_k \to 0$ cuando $k \to \infty$ (véase el ejercicio 3 de esta sección).

Una serie $\sum x_k$ es ***absolutamente convergente*** sii la serie de números reales $\sum \|x_k\|$ converge. Una serie convergente que no sea absolutamente convergente se denomina *condicionalmente* convergente. El teorema **2.9.2** y la desigualdad triangular implican el siguiente teorema:

2.9.3 Teorema *En un espacio normado completo, si $\sum x_k$ converge absolutamente, entonces $\sum x_k$ converge.*

Este teorema es útil, pues nos permite aplicar criterios de convergencia para series reales (como el criterio del cociente) a la serie $\sum \|x_k\|$ para verificar la convergencia de $\sum x_k$. Por supuesto, puede ocurrir que un criterio particular falle a pesar de que $\sum x_k$ sea convergente, en cuyo caso necesitaremos otro método.

Veamos ahora un repaso de los criterios más importantes para la *convergencia de una serie real*. Algunos de los más importantes aparecen en el siguiente teorema. En los ejercicios y, más adelante, en el capítulo 5, aparecen otros criterios de convergencia. En el apartado **vi** del teorema, se supone que el lector está familiarizado con la integración en una variable; en caso contrario, §**4.8** contiene un repaso de la teoría necesaria.

2.9.4 Teorema

i. ***Serie geométrica:*** *Si $|r| < 1$, entonces $\sum_{n=0}^{\infty} r^n$ converge a $1/(1-r)$, y diverge (no converge) si $|r| \geq 1$.*

ii. ***Criterio de comparación:*** *Si $\sum_{k=1}^{\infty} a_k$ converge, $a_k \geq 0$, y $0 \leq b_k \leq a_k$, entonces $\sum_{k=1}^{\infty} b_k$ converge; si $\sum_{k=1}^{\infty} c_k$ diverge, $c_k \geq 0$, y $0 \leq c_k \leq d_k$, entonces $\sum_{k=1}^{\infty} d_k$ diverge.*

iii. ***Criterio de la serie de Riemann:*** *$\sum_{n=1}^{\infty} 1/n^p$ converge si $p > 1$ y diverge a ∞ (es decir, las sumas parciales crecen sin límite)* si $p \leq 1$.

iv. ***Criterio del cociente:*** *Supongamos que $\lim_{n\to\infty} |a_{n+1}/a_n|$ existe y es estrictamente menor que 1. Entonces $\sum_{n=1}^{\infty} a_n$ converge absolutamente. Si el límite es estrictamente mayor que 1, entonces la serie diverge; si el límite es igual a 1, el criterio no proporciona información.*

v. ***Criterio de la raíz:*** *Supongamos que $\lim_{n\to\infty} |a_n|^{1/n}$ existe y es estrictamente menor que 1. Entonces $\sum_{n=1}^{\infty} a_n$ converge absolutamente. Si el límite es estrictamente mayor que 1, entonces la serie diverge; si el límite es igual a 1, el criterio no proporciona información.*

vi. ***Criterio de la integral:*** *Si f es continua, no negativa y monótona decreciente en $[1, +\infty[$, entonces $\sum_{n=1}^{\infty} f(n)$ y $\int_1^{\infty} f(x)dx$ son ambas convergentes o divergentes.*

vii. ***Criterio de comparación en el límite:*** *Sean $\sum_{i=1}^{\infty} a_i$ y $\sum_{i=1}^{\infty} b_i$ dos series y sea $b_i > 0$ para todo i.*

(a) si $|a_i| \leq b_i$ para todo i, o $\lim_{i\to\infty} |a_i| / b_i < \infty$, y (b) si $\sum_{i=1}^{\infty} b_i$ es convergente, entonces $\sum_{i=1}^{\infty} a_i$ es convergente.

(c) si $a_i \geq b_i$ para todo i, o $\lim_{i\to\infty} a_i / b_i > 0$, y (d) si $\sum_{i=1}^{\infty} b_i$ es divergente, entonces $\sum_{i=1}^{\infty} a_i$ es divergente.

viii. ***Series alternantes:*** *Si $\sum_{i=1}^{\infty} a_i$ es tal que los a_i alternan sus signos, son decrecientes en valor absoluto y $a_i \to 0$ cuando $i \to \infty$, entonces la serie converge. El error cometido al aproximar la suma por $s_n = \sum_{i=1}^{n} a_i$ no es mayor que $|a_{n+1}|$.*

2.9.5 Ejemplo *Sea $x_n = (1/n^2, 1/n)$ ¿Converge $\sum x_n$?*

Solución No, pues la serie armónica $\sum 1/n$ diverge, por **2.9.4iii.** ♦

2.9.6 Ejemplo *Sea $\|x_n\| \leq 1/2^n$; demuéstrese que $\sum x_n$ converge y que $\left\|\sum_{n=0}^{\infty} x_n\right\| \leq 2$.*

Solución Verifiquemos las condiciones del teorema **2.9.2.** En primer lugar,

$$\|x_k + \cdots + x_{k+p}\| \leq \|x_k\| + \cdots + \|x_{k+p}\| \leq \frac{1}{2^k} + \cdots + \frac{1}{2^{k+p}}$$
$$\leq \sum_{j=k}^{\infty} \frac{1}{2^j} = \frac{1}{2^{k-1}}$$

(por la fórmula $\sum_{n=0}^{\infty} ar^n = a/(1-r)$ para la suma de una serie geométrica). Así, dado $\varepsilon > 0$, si elegimos N tal que $1/2^{N-1} < \varepsilon$, entonces se cumplen las condiciones de **2.9.2** y podemos concluir que $\sum x_k$ converge. Además, las sumas parciales satisfacen

$$\|s_n\| = \left\| \sum_{k=0}^{n} x_k \right\| \leq \sum_{k=0}^{n} \|x_k\| \leq \sum_{k=0}^{n} \frac{1}{2^k} = 2.$$

Así, el límite s también satisface $\|s\| \leq 2$ (como en el ejemplo **2.7.7**). También podríamos probar que $\sum \|x_n\|$ converge comparándola con la serie geométrica $\sum 1/2^n$. ♦

2.9.7 Ejemplo *Analícese la convergencia de:* $\displaystyle\sum_{n=1}^{\infty} \frac{n}{3^n}$.

Solución Podemos aplicar el criterio del cociente:

$$\left| \frac{a_{n+1}}{a_n} \right| = \frac{n+1}{n} \frac{1}{3} \to \frac{1}{3} < 1,$$

de modo que la serie converge. ♦

2.9.8 Ejemplo *Determínese si* $\displaystyle\sum_{n=1}^{\infty} \frac{n}{n^2+1}$ *converge o no.*

Solución Obsérvese que para $x \geq 1$, $f(x) = x/(x^2+1)$ es positiva y continua. Como $f'(x) = (-x^2+1)/(x^2+1)^2 \leq 0$, f es monótona decreciente. Además,

$$\int_1^{\infty} \frac{x\,dx}{x^2+1} = \lim_{b\to\infty} \int_1^b \frac{x\,dx}{x^2+1}$$
$$= \lim_{b\to\infty} \left[\frac{1}{2} \log(x^2+1) \right]_1^b = \lim_{b\to\infty} \frac{1}{2} \log\left(\frac{b^2+1}{2} \right).$$

Cuando $b \to \infty$, $(1/2)\log[b^2+1]/2] \to \infty$, con lo que la serie diverge por el criterio de la integral. También podemos proceder de la siguiente manera:

$$\frac{n}{n^2+1} \geq \frac{n}{n^2+n^2} = \frac{1}{2n},$$

y comparando con la serie divergente $\frac{1}{2}\sum(1/n)$, obtenemos la divergencia de nuestra serie. ♦

2.9.9 Ejemplo *Analícese la convergencia de la serie* $\displaystyle\sum_{i=1}^{\infty} \frac{(-1)^i\sqrt{i}}{i+4}$.

Solución Sea $a_i = (-1)^i\sqrt{i}/(i+4)$. Observemos que para i grande, $|a_i|$ parece "comportarse como" $b_i = 1/\sqrt{i}$. La serie $\sum_{i=1}^{\infty} b_i$ diverge, por el criterio de las series de Riemann. Para hacer precisa la comparación entre $|a_i|$ y b_i, analizamos su cociente: $\lim_{i\to\infty}|a_i|/b_i = \lim_{i\to\infty}[i/(i+4)] = 1$, de modo que $\sum_{i=1}^{\infty}|a_i|$ también diverge; por lo tanto, *nuestra serie no es absolutamente convergente.*

La serie parece que pudiera ser alternante: los términos alternan sus signos y $\lim_{i\to\infty} a_i = 0$. Para ver si los valores absolutos $|a_i|$ forman una sucesión decreciente, es conveniente analizar la función $f(x) = \sqrt{x}/(x+4)$; su derivada es

$$f'(x) = \frac{(1/2\sqrt{x}\,)(x+4) - \sqrt{x}\cdot 1}{(x+4)^2} = \frac{2/\sqrt{x} - \sqrt{x}/2}{(x+4)^2} = \frac{4-x}{2\sqrt{x}(x+4)^2},$$

la cual es negativa para $x > 4$, por lo que $f(x)$ es decreciente para $x > 4$. Como $|a_i| = f(i)$, tendremos que $|a_4| > |a_5| > |a_6| > \cdots$, lo que implica, a su vez, que nuestra serie $\sum a_i$, omitiendo los primeros tres términos, es alternante, de donde se sigue que *la serie es convergente*; como no es absolutamente convergente, será *condicionalmente* convergente. ♦

Ejercicios de §2.9

1. Determínese si $\displaystyle\sum_{n=1}^{\infty}\left(\frac{(\operatorname{sen} n)^n}{n^2}, \frac{1}{n^2}\right)$ converge.

2. Muéstrese que la serie del ejemplo **2.9.6** converge absolutamente.

3. Sea $\sum x_k$ una serie convergente en $\mathbb{R}^n$. Muéstrese que $x_k \to 0 = (0, \ldots, 0) \in \mathbb{R}^n$.

4. Analícese la convergencia de $\sum_{n=3}^{\infty} \frac{2^n + n}{3^n - n}$.

5. Analícese la convergencia de $\sum_{n=0}^{\infty} \frac{n!}{3^n}$.

Demostraciones de los teoremas del capítulo 2

2.1.2 Proposición *En un espacio métrico, todo disco de radio ε, $D(x, \varepsilon)$, es abierto.*

Demostración Sea $y \in D(x, \varepsilon)$. Debemos encontrar ε' tal que $D(y, \varepsilon') \subset D(x, \varepsilon)$. La figura 2.1-5 sugiere probar con $\varepsilon' = \varepsilon - d(x, y)$, que es estrictamente positivo, ya que $d(x, y) < \varepsilon$. Con esta elección, que depende de y, mostraremos que $D(y, \varepsilon') \subset D(x, \varepsilon)$. Sea $z \in D(y, \varepsilon')$, entonces $d(z, y) < \varepsilon'$. Es preciso demostrar que $d(z, x) < \varepsilon$. Pero, por la desigualdad triangular, $d(z, x) \leq d(z, y) + d(y; x) < \varepsilon' + d(y, x)$, y por la elección de ε', $\varepsilon' + d(y, x) = \varepsilon$. ■

2.1.3 Proposición *En un espacio métrico (M, d),*

i. *La intersección de un número finito de subconjuntos abiertos de M es abierta.*

ii. *La unión de una colección arbitraria de subconjuntos abiertos de M es abierta.*

iii. *El conjunto vacío $\varnothing$ y el espacio total M son abiertos.*

Demostración

i. Basta con demostrar que la intersección de dos conjuntos abiertos es abierta, puesto que podemos demostrar el resultado general por inducción, escribiendo $A_1 \cap \cdots \cap A_n = (A_1 \cap \cdots \cap A_{n-1}) \cap A_n$.

Sean A y B conjuntos abiertos, y sea $C = A \cap B$; si $C = \varnothing$, C es abierto como un caso degenerado de la definición. En consecuencia, supongamos que $x \in C$. Como A y B son abiertos, existen $\varepsilon, \varepsilon' > 0$ tales que

$$D(x, \varepsilon) \subset A \quad \text{y} \quad D(x, \varepsilon') \subset B.$$

Sea ε'' el mínimo de ε y ε'. Entonces $D(x, \varepsilon'') \subset D(x, \varepsilon)$, y, por tanto, $D(x, \varepsilon'') \subset A$; de manera similar, $D(x, \varepsilon'') \subset B$, de modo que $D(x, \varepsilon'') \subset A \cap B = C$, como se pretendía.

ii. La demostración para el caso de las uniones es más fácil. Sea $U, V, \ldots$ una colección de conjuntos abiertos cuya unión es A. Si $x \in A$, $x \in U$, para algún U en la colección. Por lo tanto, como U es abierto, $D(x, \varepsilon) \subset U \subset A$ para algún $\varepsilon > 0$, lo que demuestra que A es abierto.

iii. El conjunto vacío $\varnothing$ y el espacio total M son abiertos directamente de la definición. ■

2.4.2 Teorema *Un conjunto $A \subset M$ es cerrado sii todos los puntos de acumulación de A pertenecen a A.*

Demostración En primer lugar, supongamos que A es cerrado. Entonces, $M \backslash A$ es abierto. Así, si $x \in M \backslash A$, existe $\varepsilon > 0$ tal que $D(x, \varepsilon) \subset M \backslash A$; es decir, $D(x, \varepsilon) \cap A = \varnothing$. Así, x no es un punto de acumulación y entonces A contiene a todos sus puntos de acumulación. Recíprocamente, supongamos que A contiene a todos sus puntos de acumulación. Sea $x \in M \backslash A$. Como x no es punto de acumulación de A y $x \notin A$, existe $\varepsilon > 0$ tal que $D(x, \varepsilon) \cap A = \varnothing$; es decir, $D(x, \varepsilon) \subset M \backslash A$. Por lo tanto, $M \backslash A$ es abierto y entonces A es cerrado. ■

2.5.2 Proposición *Si $A \subset M$, $\mathrm{cl}(A)$ es el conjunto formado por A junto con los puntos de acumulación de A; es decir, $\mathrm{cl}(A) = A \cup \{$puntos de acumulación de $A\}$.*

Demostración Sea $B = A \cup \{x \mid x$ es un punto de acumulación de $A\}$. Por **2.4.2**, cualquier conjunto cerrado que contiene a A también debe contener a B. Si B es cerrado, entonces será el menor conjunto cerrado que contiene a A, con lo que $B = \mathrm{cl}(A)$. Para demostrar que B es cerrado, usamos **2.4.2**; sea y un punto de acumulación de B. Si $\varepsilon > 0$, entonces $D(y, \varepsilon)$ contiene otros puntos de B. Si z es uno de tales puntos, entonces $z \in A$ o z es un punto de acumulación de A. En este último caso, $D(z, \varepsilon - d(z, y))$ es un conjunto abierto que contiene a z y, entonces, **2.4.1** implica que este conjunto contiene otros puntos de A, necesariamente distintos de y. Así, y es un punto de acumulación de A y entonces $y \in B$, por lo que B es cerrado. ■

2.6.2 Proposición *Sea $A \subset M$. Entonces $x \in \partial A$ sii para todo $\varepsilon > 0$, $D(x, \varepsilon)$ contiene puntos de A y de $M \backslash A$. (Estos puntos pueden incluir al propio x.)*

Demostración Sea $x \in \partial A = \mathrm{cl}(A) \cap \mathrm{cl}(M \backslash A)$. Entonces, o bien $x \in A$, o bien $x \in M \backslash A$. En el primer caso, la proposición **2.5.2** implica que x es punto de acumulación de $M \backslash A$ y se cumple el resultado. El caso $x \in M \backslash A$ y el recíproco son similares. ■

2.7.2 Proposición *Una sucesión x_k en M converge a $x \in M$ sii para todo $\varepsilon > 0$ existe N tal que $k \geq N$ implica $d(x, x_k) < \varepsilon$.*

Demostración Supongamos que $x_k \to x$ y $\varepsilon > 0$. Como $D(x, \varepsilon)$ es abierto, existe N tal que $k \geq N$ implica $x_k \in D(x, \varepsilon)$ o $d(x, x_k) < \varepsilon$, como se pedía. Recíprocamente, supongamos que la condición es válida y U es una vecindad de x. Sea $\varepsilon > 0$ tal que $D(x, \varepsilon) \subset U$; entonces, existe N tal que $k \geq N$ implica $d(x, x_k) < \varepsilon$; es decir, $x_k \in D(x, \varepsilon) \subset U$, por lo que $x_k \to x$, por definición. ■

2.7.3 Proposición *Sean v_k y w_k sucesiones de vectores en un espacio normado (como por ejemplo $\mathbb{R}^n$), λ_k una sucesión de números en $\mathbb{R}$, $\lambda \in \mathbb{R}$ una constante y u un vector constante. Si $v_k \to v$, $w_k \to w$ y $\lambda_k \to \lambda$, entonces*

i. $v_k + w_k \to v + w$.

ii. $\lambda v_k \to \lambda v$.

iii. $\lambda_k u \to \lambda u$.

iv. $\lambda_k v_k \to \lambda v$.

v. *Si $\lambda_k \neq 0$ y $\lambda \neq 0$, entonces* $\dfrac{1}{\lambda_k} v_k \to \dfrac{1}{\lambda} v$.

Demostración En cada caso la demostración es similar a la de **1.2.7**, por lo que dejamos que el lector la adapte por sí mismo. Para ahorrar esfuerzo, obsérvese que, como en **1.2.7, iv** implica **ii** y **iii**, por lo que sólo hay que probar **i, iv** y **v**. ■

2.7.4 Proposición *$v_k \to v$ en $\mathbb{R}^n$ si y sólo si cada sucesión de coordenadas converge a la coordenada correspondiente de v como una sucesión en $\mathbb{R}$. Es decir, $\lim_{k\to\infty} v_k = v$ en $\mathbb{R}^n$ sii $\lim_{k\to\infty} v_k^i = v^i$ en $\mathbb{R}$ para cada $i = 1, 2, \ldots, n$.*

Demostración Si $\delta(v, v_k) = \max\{|v^1 - v_k^1|, \ldots, |v^n - v_k^n|\}$ entonces $\delta(v, v_k) \leq \|v - v_k\| \leq \sqrt{n}\,\delta(v, v_k)$, por **1.6.7**. Si $v_k \to v$ en $\mathbb{R}^n$, $1 \leq j \leq n$ y $\varepsilon > 0$, entonces existe un entero K tal que $\|v - v_k\| < \varepsilon$ siempre que $k \geq K$. En consecuencia, para ese k, tenemos que $|v^j - v_k^j| \leq \delta(v, v_k) \leq \|v - v_k\| < \varepsilon$, por lo que $v_k^j \to v^j$ en $\mathbb{R}$. La demostración de la recíproca muestra la importancia de que n sea finito. Supongamos que para cada $j = 1, 2, \ldots, n$, $\lim_{k\to\infty} v_k^j = v^j$ y sea $\varepsilon > 0$. Para todo $\varepsilon_0 > 0$ existen enteros $K_1, K_2, K_3, \ldots, K_n$ tales que

$$
\begin{aligned}
|v^1 - v_k^1| &< \varepsilon_0 \quad \text{si} \quad k \geq K_1,\\
|v^2 - v_k^2| &< \varepsilon_0 \quad \text{si} \quad k \geq K_2,\\
&\ldots \qquad\qquad \ldots\\
|v^n - v_k^n| &< \varepsilon_0 \quad \text{si} \quad k \geq K_n.
\end{aligned}
$$

Sólo hay un número finito de K_i, luego uno de ellos es el máximo. Sea $K = \text{máx}(K_1, K_2, \ldots, K_n)$. Si $k \geq K$, **1.6.7** implica de nuevo que $\|v - v_k\| \leq \sqrt{n}\,\delta(v, v_k) \leq \sqrt{n}\varepsilon_0$. Si hacemos esto con $\varepsilon_0 = \varepsilon/\sqrt{n}$, obtenemos $\|v - v_k\| < \varepsilon$ cuando $k \geq K$, así que $v_k \to v$ en $\mathbb{R}^n$. ■

2.7.6 Proposición

i. *Un conjunto $A \subset M$ es cerrado sii para cada sucesión $x_k \in A$ que converge en M, el límite es un elemento de A.*

ii. *Para un conjunto $B \subset M$, $x \in \text{cl}(B)$ sii existe una sucesión $x_k \in B$ tal que $x_k \to x$.*

Demostración

i. En primer lugar, supongamos que A es cerrado, y que $x_k \to x$. Entonces x es un punto de acumulación de A, ya que cualquier vecindad de x contiene a $x_k \in A$ para k suficientemente grande. El teorema **2.4.2** implica entonces que $x \in A$.

Recíprocamente, usaremos el teorema **2.4.2** para demostrar que A es cerrado. Sean x un punto de acumulación de A y elijamos $x_k \in D(x, 1/k) \cap A$. Entonces $x_k \to x$, ya que para cada $\varepsilon > 0$ podemos elegir $N \geq 1/\varepsilon$, con lo que $k \geq N$ implica $x_k \in D(x, \varepsilon)$. (Véase la figura 2.P-1.) Por hipótesis $x \in A$, luego A es cerrado.

ii. El argumento en este caso es similar y lo dejaremos como ejercicio. ■

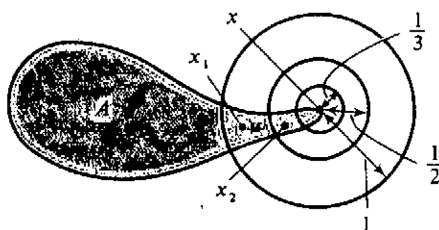


FIGURA 2.P-1 Puntos de acumulación de un conjunto

2.8.3 Proposición *Una sucesión convergente en un espacio normado o métrico está acotada.*

Demostración Seguiremos el procedimiento de **1.2.6.** Si $x_k \to x$, existe N tal que $d(x_n, x) < 1$ cuando $n \geq N$. Así, $x_n \in D(x, 1)$ si $n \geq N$. Sea

$$R = \text{máx}\{1, d(x_1, x), \ldots, d(x_{N-1}, x)\}.$$

Entonces $d(x, x_n) \leq R$ para todo n, con lo que $x_n \in D(x, R)$ para todo n. ■

2.8.4 Proposición

i. *Toda sucesión convergente en un espacio métrico es una sucesión de Cauchy.*

ii. *Una sucesión de Cauchy en un espacio métrico debe estar acotada.*

iii. *Si una subsucesión de una sucesión de Cauchy converge a x, entonces la sucesión converge a x.*

Demostración La demostración de **i** es como la de **1.4.2,** la de **ii** como la de **1.4.6,** y la de **iii** como la de **1.4.7.** En cada caso reemplazamos los valores absolutos por las distancias correspondientes; por ejemplo, léase $d(x, x_m)$ en lugar de $|x - x_m|$. ■

2.8.5 Teorema *Una sucesión x_k en $\mathbb{R}^n$ converge a un punto en $\mathbb{R}^n$ sii es una sucesión de Cauchy.*

Demostración Si x_k converge a x, entonces para cada $\varepsilon > 0$, elegimos N tal que $k \geq N$ implique $\|x_k - x\| < \varepsilon/2$. En consecuencia, si $k, l \geq N$, $\|x_k - x_l\| = \|(x_k - x) + (x - x_l)\| \leq \|x_k - x\| + \|x - x_l\| < \varepsilon/2 + < \varepsilon/2 = \varepsilon$, por la desigualdad triangular, luego x_k es una sucesión de Cauchy.

Recíprocamente, sea x_k una sucesión de Cauchy. Como $|x_k^i - x_l^i| \leq \|x_k - x_l\|$, las componentes también son sucesiones de Cauchy en la recta real. Por la completitud de $\mathbb{R}$ y el teorema **1.4.4,** x_k^i converge, digamos, a x^i. La proposición **2.7.4** implica que x_k converge a $x = (x^1, \ldots, x^n)$. ■

2.8.7 Proposición *Si x_k es una sucesión en un espacio métrico M y $x \in M$, entonces*

i. *x es un punto límite sii para cada $\varepsilon > 0$ y cada entero N existe $k > N$ tal que $d(x_k, x) < \varepsilon$.*

ii. *x es un punto límite sii existe una subsucesión convergente a x.*

iii. *$x_k \to x$ sii toda subsucesión converge a x.*

iv. $x_k \to x$ *sii toda subsucesión de* x_k *tiene a su vez una subsucesión que converge a* x.

Demostración Esta demostración es similar a la de **1.5.2**, así que dejamos que el lector la adapte a este caso. ■

2.9.2 Teorema *Sea* V *un espacio normado completo (como por ejemplo* $\mathbb{R}^n$*). Una serie* $\sum x_k$ *en* V *converge sii para todo* $\varepsilon > 0$ *existe* N *tal que* $k \geq N$ *implica*

$$\|x_k + x_{k+1} + \cdots + x_{k+p}\| < \varepsilon \text{ para todo entero } p = 0, 1, 2, \ldots.$$

Demostración Sea $s_k = \sum_{i=1}^{k} x_i$. Por la completitud, $\sum x_k$ converge sii s_k es una sucesión de Cauchy. Esto es cierto sii para cada $\varepsilon > 0$ existe N tal que $l \geq N$ implica $\|s_l - s_{l+q}\| < \varepsilon$ para todo $q = 1, 2, \ldots$. Pero $\|s_{l+q} - s_l\| = \|x_{l+1} + \cdots + x_{l+q}\|$ por lo que el resultado se sigue tomando $k = l + 1$ y $p = q - 1 = 0, 1, 2, \ldots$. ■

2.9.3 Teorema *En un espacio normado completo,* $\sum x_k$ *converge absolutamente, entonces* $\sum x_k$ *converge.*

Demostración Se sigue del teorema **2.9.2** y de la desigualdad triangular:

$$\|x_k + \cdots + x_{k+p}\| \leq \|x_k\| + \cdots + \|x_{k+p}\|. \quad ■$$

2.9.4 Teorema

i. ***Serie geométrica:*** *Si* $|r| < 1$, *entonces* $\sum_{n=0}^{\infty} r^n$ *converge a* $1/(1-r)$, *y diverge (no converge) si* $|r| \geq 1$.

ii. ***Criterio de comparación:*** *Si* $\sum_{k=1}^{\infty} a_k$ *converge,* $a_k \geq 0$ *y* $0 \leq b_k \leq a_k$, *entonces* $\sum_{k=1}^{\infty} b_k$ *converge; si* $\sum_{k=1}^{\infty} c_k$ *diverge,* $c_k \geq 0$ *y* $0 \leq c_k \leq d_k$, *entonces* $\sum_{k=1}^{\infty} d_k$ *diverge.*

iii. ***Criterio de la serie de Riemann:*** $\sum_{n=1}^{\infty} 1/n^p$ *converge si* $p > 1$ *y diverge a* ∞, *(es decir, las sumas parciales crecen sin límite)* si $p \leq 1$.

iv. ***Criterio del cociente:*** *Supongamos que* $\lim_{n\to\infty} |a_{n+1}/a_n|$ *existe y es estrictamente menor que* 1. *Entonces* $\sum_{n=1}^{\infty} a_n$ *converge absolutamente. Si el límite es estrictamente mayor que* 1, *entonces la serie diverge; si el límite es igual a* 1, *el criterio no proporciona información.*

v. ***Criterio de la raíz:*** *Supongamos que* $\lim_{n\to\infty}|a_n|^{1/n}$ *existe y es estrictamente menor que* 1. *Entonces* $\sum_{n=1}^{\infty} a_n$ *converge absolutamente. Si el límite es estrictamente mayor que* 1, *entonces la serie diverge; si el límite es igual a* 1, *el criterio no proporciona información.*

vi. ***Criterio de la integral:*** *Si f es continua, no negativa y monótona decreciente en* $[1, +\infty[$, *entonces* $\sum_{n=1}^{\infty} f(n)$ *y* $\int_1^{\infty} f(x)dx$ *son ambas convergentes o divergentes.*

vii. ***Criterio de comparación en el límite:*** *Sean* $\sum_{i=1}^{\infty} a_i$ *y* $\sum_{i=1}^{\infty} b_i$ *dos series y sea* $b_i >$ 0 *para todo i.*

Si (a) $|a_i| \le b_i$ *para todo i, o* $\lim_{i\to\infty}|a_i|/b_i < \infty$ *y si (b)* $\sum_{i=1}^{\infty} b_i$ *es convergente, entonces* $\sum_{i=1}^{\infty} a_i$ *es convergente.*

Si (c) $a_i \ge b_i$ *para todo i, o* $\lim_{i\to\infty} a_i/b_i > 0$ *y si (d)* $\sum_{i=1}^{\infty} b_i$ *es divergente, entonces* $\sum_{i=1}^{\infty} a_i$ *es divergente.*

viii. ***Series alternantes:*** *Si* $\sum_{i=1}^{\infty} a_i$ *es tal que los* a_i *alternan sus signos, son decrecientes en valor absoluto y* $a_i \to 0$ *cuando* $i \to \infty$, *entonces la serie converge. El error cometido al aproximar la suma por* $s_n = \sum_{i=1}^{n} a_i$ *no es mayor que* $|a_{n+1}|$.

Demostración

i. Usaremos el siguiente

Lema de la potencia

$$\lim_{n\to\infty} r^n = \begin{cases} \infty & \text{si } r > 1 \\ 1 & \text{si } r = 1 \\ 0 & \text{si } 0 \le r < 1. \end{cases}$$

Demostración En primer lugar, consideremos el caso $r > 1$. Escribimos r como $1 + s$, donde $s > 0$. Si desarrollamos $r^n = (1 + s)^n$, obtenemos $r^n = 1 + ns +$ (otros términos positivos). En consecuencia, $r^n \ge 1 + ns$, que tiende a ∞ si $n \to \infty$. En segundo lugar, si $r = 1$, entonces $r^n = 1$ para todo n, por lo que $\lim_{n\to\infty} r^n = 1$. Por último, si $0 \le r < 1$ y excluimos el sencillo caso $\rho = 0$, sea $\rho = 1/r$, de modo que $\rho > 1$ y, por lo tanto, $\lim_{n\to\infty} \rho^n = \infty$. En consecuencia, $\lim_{n\to\infty} r^n = \lim_{n\to\infty} 1/\rho^n = 0$. ▼

Por un cálculo algebraico elemental,

$$1 + r + r^2 + \cdots + r^n = \frac{1 - r^{n+1}}{1 - r}$$

si $r \ne 1$. El lema de la potencia implica que $r^{n+1} \to 0$ cuando $n \to \infty$ si $|r| < 1$, y que $|r|^{n+1} \to \infty$ si $|r| > 1$. Así, tenemos la convergencia si $|r| < 1$ y la divergencia si $|r| > 1$. Obviamente $\sum_{n=0}^{\infty} r^n$ diverge si $|r| = 1$, pues $r^n \not\to 0$.

ii. Las sumas parciales de la serie $\sum_{k=1}^{\infty} a_k$ forman una sucesión de Cauchy, por lo que las sumas parciales de la serie $\sum_{k=1}^{\infty} b_k$ también forman una sucesión de Cauchy, ya que para todo k y p tenemos $b_k + b_{k+1} + \cdots + b_{k+p} \leq a_k + a_{k+1} + \cdots + a_{k+p}$. Por lo tanto $\sum_{k=1}^{\infty} b_k$ converge. Una serie positiva sólo puede diverger a $+\infty$, así que, dado $M > 0$, podemos encontrar k_0 tal que $k \geq k_0$ implique $c_1 + c_2 + \cdots + c_k \geq M$. En consecuencia, si $k \geq k_0$, $d_1 + d_2 + \cdots + d_k \geq M$, de modo que $\sum_{k=1}^{\infty} d_k$ también diverge a ∞.

iii. Supongamos en primer lugar que $p \leq 1$; en este caso, $1/n^p \geq 1/n$ para todo $n = 1, 2, \ldots$. En consecuencia, por **ii**, $\sum_{n=1}^{\infty}(1/n^p)$ diverge si la serie armónica $\sum_{n=1}^{\infty}(1/n)$ diverge; pero ya hemos demostrado esto en **1.2.20.**

Supongamos ahora que $p > 1$. Sea

$$s_k = \frac{1}{1^p} + \frac{1}{2^p} + \frac{1}{3^p} + \cdots + \frac{1}{k^p},$$

entonces s_k es una sucesión creciente de números reales positivos. Por otro lado,

$$\begin{aligned} s_{2^k - 1} &= \frac{1}{1^p} + \left(\frac{1}{2^p} + \frac{1}{3^p}\right) + \left(\frac{1}{4^p} + \frac{1}{5^p} + \frac{1}{6^p} + \frac{1}{7^p}\right) \\ &\quad + \cdots + \left(\frac{1}{(2^{k-1})^p} + \cdots + \frac{1}{(2^k - 1)^p}\right) \\ &\leq \frac{1}{1^p} + \frac{2}{2^p} + \frac{4}{4^p} + \cdots + \frac{2^{k-1}}{(2^{k-1})^p} \\ &= \frac{1}{1^{p-1}} + \frac{1}{2^{p-1}} + \frac{1}{4^{p-1}} + \cdots + \frac{1}{(2p-1)^{p-1}} \\ &< \frac{1}{1 - \frac{1}{2^{p-1}}} \end{aligned}$$

(¿por qué?). Así, la sucesión $\{s_k\}$ está acotada superiormente por $\left(1 - \frac{1}{2^{p-1}}\right)^{-1}$; por lo tanto, $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converge.

Nota. También podemos demostrar **iii** mediante el criterio de la integral para series positivas (véase el apartado **vi** del teorema). Sin embargo, la demostración que hemos dado también prueba el *criterio de condensación de Cauchy:* Sea $\sum a_n$ una serie de términos positivos tales que $a_{n+1} \leq a_n$. Entonces $\sum a_n$ converge sii $\sum_{j=1}^{\infty} 2^j a_{2^j}$ converge (G. J. Porter, "An Alternative to the Integral Test for Infinite Series", *Am. Math. Monthly* **79** (1972), pág. 634).

iv. Supongamos que $\lim_{n\to\infty}|a_{n+1}/a_n| = r < 1$. Elijamos r' tal que $r < r' < 1$ y N tal que $n \geq N$ implique $|a_{n+1}/a_n| < r'$. Por inducción, $|a_{N+p}| < |a_N|(r')^p$. Consideremos la serie $|a_1| + \cdots + |a_N| + |a_N|r' + |a_N|(r')^2 + |a_N|(r')^3 + \cdots$, que converge a

$$|a_1| + \cdots + |a_{N-1}| + \frac{|a_N|}{1-r'}.$$

Por **ii**, podemos concluir que $\sum_{k=1}^{\infty}|a_k|$ converge. Si $\lim_{n\to\infty}|a_{n+1}/a_n| = r > 1$, elijamos r' tal que $1 < r' < r$ y N tal que $n \geq N$ implique que $|a_{n+1}/a_n| > r'$. Entonces, $|a_{N+p}| > (r')^p|a_N|$, luego $\lim_{n\to\infty}|a_N| = \infty$, mientras que el límite debería ser cero si la serie convergiera. Así pues, $\sum_{k=1}^{\infty} a_k$ diverge. Para ver que el criterio no sirve si $\lim_{n\to\infty}|a_{n+1}/a_n| = 1$, consideremos las series $1 + 1 + 1 + \cdots$, y $\sum_{n=1}^{\infty} 1/n^p$ para $p > 1$. En ambos casos, $\lim_{n\to\infty}|a_{n+1}/a_n| = 1$, pero mientras que la primera serie diverge, la segunda converge.

v. Supongamos que $\lim_{n\to\infty}(|a_n|)^{1/n} = r < 1$. Elijamos r' tal que $r < r' < 1$ y N tal que $n \geq N$ implique $|a_n|^{1/n} < r'$; en otras palabras, $|a_n| < (r')^n$. La serie $|a_1| + |a_2| + \cdots + |a_{N-1}| + (r')^N + (r')^{N+1} + \cdots$ converge a $|a_1| + |a_2| + \cdots + |a_{N-1}| + (r')^n/(1 - r')$; **ii** implica entonces que $\sum_{k=1}^{\infty} a_k$ converge. Si $\lim_{n\to\infty}(|a_n|)^{1/n} = r > 1$, elijamos r' tal que $1 < r' < r$ y N tal que $n \geq N$ implique que $|a_n|^{1/n} > r'$, o, en otras palabras, $|a_n| > (r')^n$. En consecuencia, $\lim_{n\to\infty}|a_n| = \infty$ y, por tanto, $\sum_{k=1}^{\infty} a_k$ diverge.

Para mostrar que el criterio no sirve si $\lim_{n\to\infty}(|a_n|)^{1/n} = 1$, observemos que, del cálculo elemental,

$$\lim_{n\to\infty}\left(\frac{1}{n}\right)^{1/n} = 1 \quad \text{y} \quad \lim_{n\to\infty}\left(\frac{1}{n^2}\right)^{1/n} = 1$$

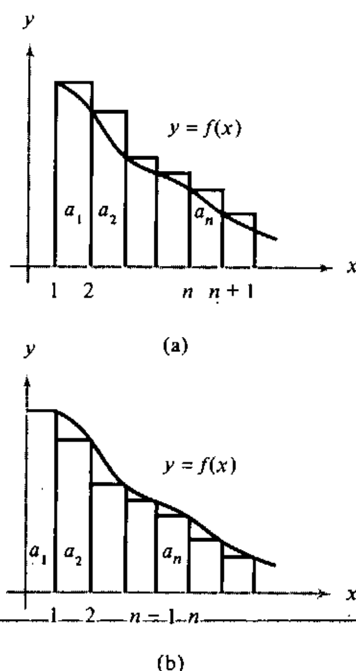
(se toman logaritmos y se utiliza el hecho de que $(\log x)/x \to 0$ si $x \to \infty$). Pero $\sum_{n=1}^{\infty} 1/n$ diverge y $\sum_{n=1}^{\infty} 1/n^2$ converge.

vi. Para este apartado aceptaremos algunos hechos elementales del cálculo relativos a las integrales (consúltese §**4.8** como repaso). En la figura 2.P-2a, los rectángulos de áreas $a_1, a_2, \ldots, a_n$ encierran más área que la que hay por debajo de la curva, entre $x = 1$ y $x = n + 1$. En consecuencia,

$$a_1 + a_2 + \cdots + a_n \geq \int_1^{n+1} f(x)dx.$$

Si ahora analizamos la figura 2.P-2b y tomamos el área entre $x = 1$ y $x = n$, obtenemos

$$a_2 + a_3 + \cdots + a_n \leq \int_1^{n} f(x)dx.$$



(b)

FIGURA 2.P-2 Desigualdades necesarias para el criterio de la integral

Sumamos a_1 a ambos miembros para obtener

$$a_1 + a_2 + a_3 + \cdots + a_n \leq a_1 + \int_1^n f(x)dx.$$

Y combinando los dos resultados llegamos a

$$\int_1^{n+1} f(x)dx \leq a_1 + a_2 + \cdots + a_n \leq a_1 + \int_1^n f(x)dx.$$

Si la integral $\int_1^\infty f(x)dx$ es finita, la desigualdad de la derecha implica que la serie $\sum_{n=1}^\infty a_n$ también es finita, por la propiedad de completitud de $\mathbb{R}$. Pero si $\int_1^\infty f(x)dx$ es infinita, la desigualdad de la izquierda muestra que la serie también es infinita. Por lo tanto, la serie y la integral son ambas convergentes o divergentes.

vii. Supongamos, por ejemplo, que $\lim_{i\to\infty} |a_i|/b_i = M < \infty$ y que $b_i > 0$ para todo i. Si i es suficientemente grande, tendremos $|a_i|/b_i < M + 1$; es decir, $|a_i| < (M + 1)b_i$. Si $\sum b_i$ converge, también lo hace $\sum(M + 1)b_i$, y el criterio de comparación implica que $\sum a_i$ converge. Los demás casos son análogos.

viii. Sea $\sum_{n=1}^{\infty} a_i$ una serie alternante. Si $a_1 > 0$ y definimos $b_i = (-1)^{i+1}a_i$, entonces los b_i son todos positivos y la serie es igual a $b_1 - b_2 + b_3 - b_4 + b_5 \cdots$. Además, tenemos que $b_1 > b_2 > b_3 > \cdots$ y que $\lim_{i\to\infty} b_i = 0$. Cada suma parcial par S_{2n} se puede agrupar como $(b_1 - b_2) + (b_3 - b_4) + \cdots + (b_{n-1} - b_n)$, que es una serie de términos positivos, por lo que $S_2 \leq S_4 \leq S_6 \leq \cdots$. Las sumas parciales impares S_{2n+1} se pueden agrupar como $b_1 - (b_2 - b_3) - (b_4 - b_5) - \cdots - (b_{2n} - b_{2n+1})$, que es una suma de términos negativos (excepto el primero), de modo que $S_1 \geq S_3 \geq S_5 \geq \cdots$. Obsérvese ahora que $S_{2n+1} = S_{2n} + b_{2n+1} \geq S_{2n}$, luego las sumas parciales pares S_{2n} forman una sucesión creciente acotada superiormente por cualquier elemento de la sucesión decreciente de sumas parciales impares. Por la propiedad de la sucesión monótona, la sucesión S_{2n} tiende a un límite S_{par}. De forma análoga, la sucesión decreciente S_{2n+1} tiende a un límite S_{impar}.

Así pues, tenemos que $S_2 \leq S_4 \leq S_6 \leq \cdots \leq S_{2n} \leq \cdots \leq S_{\text{par}} \leq S_{\text{impar}} \leq \cdots \leq S_{2n+1} \leq \cdots \leq S_3 \leq S_1$. Como $S_{2n+1} - S_{2n} = a_{2n+1}$, que tiende a cero si $n \to \infty$, y la diferencia $S_{\text{par}} - S_{\text{impar}}$ es menor que $S_{2n+1} - S_{2n}$, ésta debe anularse; es decir, $S_{\text{impar}} = S_{\text{par}}$. Llamemos a este valor común S. Entonces

$$|S_{2n} - S| \leq |S_{2n} - S_{2n+1}| = b_{2n+1} = |a_{2n+1}|$$

y

$$|S_{2n+1} - S| \leq |S_{2n+1} - S_{2n+2}| = b_{2n+2} = |a_{2n+2}|,$$

por lo cual, cada diferencia $|S_n - S|$ es menor que $|a_{n+1}|$. Como $a_{n+1} \to 0$, obtenemos que $S_n \to S$ si $n \to \infty$. Este argumento muestra también que cualquier cola de una serie alternante no es mayor que el primer término omitido de la suma parcial. ■

Ejemplos resueltos del capítulo 2

Ejemplo 2.1 *Sea $S = \{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| \leq 1, |x_2| < 1\}$. ¿Es S abierto, cerrado o ninguna de las dos cosas? ¿Cuál es el interior de S?*

Solución S no es abierto, pues no existe ninguna vecindad en torno a los puntos de S con $x_1 = 1$ que esté totalmente contenida en S (véase la figura 2.E-1). Por otro lado, S no es cerrado, pues

$$\mathbb{R}^2 \setminus S = \{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| > |x_2| \geq 1\}$$

y no existe ninguna vecindad en torno a ningún punto de $\mathbb{R}^2 \setminus S$ con $x_2 = 1$ que esté contenida en $\mathbb{R}^2 \setminus S$.

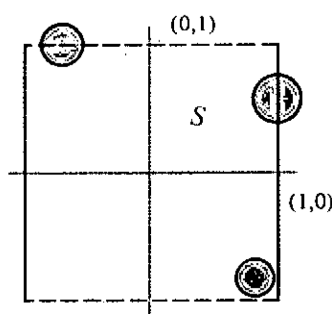


FIGURA 2.E-1 Determínense las propiedades topológicas de este conjunto

También podemos ver que S no es cerrado si observamos que la sucesión $(0, 1 - 1/n)$ converge pero el límite $(0, 1)$ no está en S (véase la proposición **2.7.6**).

Afirmamos que $\text{int}(S) = \{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| < 1, |x_2| < 1\}$, lo que verificamos mostrando que los elementos de este conjunto son los puntos interiores de S. Si $|x_1| < 1$ y $|x_2| < 1$, entonces el disco con centro en (x_1, x_2) y radio $r = \text{mínimo } \{1 - |x_1|, 1 - |x_2|\}$ está en S. Como ya hemos visto, los demás puntos de S no son puntos interiores. Una vez que el estudiante se familiarice con este tipo de argumentos, se pueden omitir algunos detalles. ♦

Ejemplo 2.2 *Muéstrese que si x es un punto de acumulación de un conjunto $S \subset \mathbb{R}^n$, entonces todo conjunto abierto que contenga a x contiene a su vez infinitos puntos de S.*

Solución Usaremos el método de reducción al absurdo. Supongamos que existe un conjunto abierto U que contiene a x pero solamente un número finito de puntos de S. Sean $x_1, x_2, \ldots, x_m$ los puntos de S en U distintos de x. Sea ε el mínimo de los números $d(x, x_1), d(x, x_2), \ldots, d(x, x_m)$, de modo que $\varepsilon > 0$. Entonces $D(x, \varepsilon)$ no contiene puntos de S distintos de x, lo que contradice el hecho de que x es un punto de acumulación de S. El lector debería proporcionar también una demostración *directa* de este resultado. ♦

Ejemplo 2.3 *Si $x = \sup(S)$ para $S \subset \mathbb{R}$, muéstrese que $x \in \text{cl}(S)$.*

Solución Por la proposición **2.5.2**, basta mostrar que, o bien $x \in S$, o bien x es un punto de acumulación de S. La proposición **1.3.2** implica que para cualquier $\varepsilon > 0$

existe $y \in S$ tal que $d(x, y) < \varepsilon$. Esto significa que si $x \notin S$, entonces x es un punto de acumulación de S. ♦

Ejemplo 2.4 *Muéstrese que una sucesión de un espacio métrico puede converger, cuando mucho a un punto (los límites son únicos).*

Solución Supongamos que $x_k \to x$ y $x_k \to y$. Dado $\varepsilon > 0$, sea N tal que $k \geq N$ implique que $d(x_k, x) < \varepsilon/2$ y sea M tal que $k \geq M$ implique que $d(x_k, y) < \varepsilon/2$. Si $k \geq N$ y $k \geq M$, entonces $d(x, y) \leq d(x, x_k) + d(x_k, y) < \varepsilon$ (por la desigualdad triangular). Como $0 \leq d(x, y) < \varepsilon$ se cumple para cada $\varepsilon > 0$, $d(x, y) = 0$, por lo que $x = y$. ♦

Ejemplo 2.5 ***Notación "O mayúscula" y "o minúscula"*** *Escribimos $f = O(g)$ si $g(x) > 0$ para $x \in \mathbb{R}$ suficientemente grande y $f(x)/g(x)$ está acotado para x suficientemente grande. Escribimos $f = o(g)$ si f/g tiende a cero cuando x tiende a $+\infty$. También escribimos $f \sim g$ (que se lee f es **asintótica a** g) si $f/g \to 1$ cuando $x \to \infty$. Demuéstrese lo siguiente:*

1. $x^2 + x = O(x^2)$.

2. $x^2 + x \sim x^2$.

3. $\exp(\sqrt{\log x}) = o(x)$.

Solución Obsérvese que si f es asintótica a g, esto implica de inmediato que $f = O(g)$ (¿por qué?). Así, **1** se sigue de **2**. Pero **2** es fácil, pues sabemos que $(x^2 + x)/x^2 = 1 + 1/x$ tiende a 1 cuando x tiende a infinito. Para demostrar **3**, observamos que $\exp(\log x) = x$, de modo que $(\exp\sqrt{\log x})/x = \exp(\sqrt{\log x} - \log x$. Puesto que $\log x \to \infty$ cuando $x \to \infty$, $(\sqrt{\log x})/\log x \to 0$ cuando $x \to \infty$, de modo que, para x suficientemente grande, $\sqrt{\log x} \leq (\log x)/2$ y, por lo tanto, para x suficientemente grande,

$$\frac{\exp\sqrt{\log x}}{x} \leq \frac{1}{x}\left(\exp\left(\frac{\log x}{2}\right)\right) = \frac{1}{\sqrt{x}},$$

que tiende a cero cuando $x \to \infty$. ♦

Ejemplo 2.6 *Recuérdese que podemos definir e^x como*

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots.$$

(Por el criterio del cociente, esta serie converge para todo $x \in \mathbb{R}$. Por lo tanto, esta definición de e^x tiene sentido.) Demuéstrese que $e = e^1$ es un número irracional.

Solución Supongamos que $e = a/b$, con a y b enteros. Sea k un entero $k > b$ y tomemos $\alpha = k!(e - 2 - 1/2! - 1/3! - \cdots - 1/k!)$, con lo que α es también un entero distinto de cero. Como $e = 2 + 1/2! + 1/3! + \cdots$,

$$\alpha = \frac{1}{k+1} + \frac{1}{(k+1)(k+2)} + \cdots \leq \frac{1}{k+1} + \frac{1}{(k+1)^2} + \cdots = \frac{1}{k}.$$

(La última igualdad se obtiene mediante la fórmula para la suma de una serie geométrica $y + y^2 + \cdots = y/(1-y)$, $0 \leq y < 1$.) Pero $\alpha < 1/k$ es imposible si α es un entero distinto de cero. Así pues, también $e = a/b$ es imposible, por lo que e es irracional. ♦

Nota. Sorprendentemente, la demostración de que e^r es irracional para r racional no es nada sencilla, y la demostración de que π es irracional es aún más difícil. Véase, por ejemplo, G.H. Hardy y E.M. Wright, *An Introduction to the Theory of Numbers*, 4ª ed., Oxford University Press, Nueva York, 1960. De hecho, e y π son ***números trascendentes***, lo que significa que no son raíces de ningún polinomio con coeficientes racionales. Este hecho fue descubierto por Charles Hermite y C.L.F. Lindemann en 1873 y 1882. Para una descripción elemental, véase M. Spivak, *Calculus*, W.A. Benjamin Co., Menlo Park, California.[2]

Ejercicios del capítulo 2

1. Analícese si los siguientes conjuntos son abiertos o cerrados:

a. $]1, 2[$ en $\mathbb{R}^1 = \mathbb{R}$

b. $[2, 3]$ en $\mathbb{R}$

c. $\cap_{n=1}^{\infty} [-1, 1/n[$ en $\mathbb{R}$

d. $\mathbb{R}^n$ en $\mathbb{R}^n$

e. Un hiperplano en $\mathbb{R}^n$

f. $\{r \in]0, 1[\mid r$ es racional$\}$ en $\mathbb{R}$

g. $\{(x, y) \in \mathbb{R}^2 \mid 0 < x \leq 1\}$ en $\mathbb{R}^2$

h. $\{x \in \mathbb{R}^n \mid \|x\| = 1\}$ en $\mathbb{R}^n$

2. Determínense el interior, la clausura y la frontera de cada uno de los conjuntos del ejercicio 1.

3. Sea U abierto en M y $U \subset A$. Muéstrese que $U \subset \text{int}(A)$. ¿Cuál es la proposición correspondiente para conjuntos cerrados?

[2] Traducción al castellano: M. Spivak, *Calculus*, Reverté, Barcelona, 1994. (*N. del R.T.*)

4. **a.** Muéstrese que si $x_n \to x$ en un espacio métrico M, entonces $x \in \mathrm{cl}\{x_1, x_2, \ldots\}$. ¿En qué caso es x un punto de acumulación?

b. ¿Puede una sucesión tener más de un punto de acumulación?

c. Si x es un punto de acumulación de un conjunto A, demuéstrese que existe una sucesión de puntos *distintos* de A que converge a x.

5. Muéstrese que $x \in \mathrm{int}(A)$ sii existe $\varepsilon > 0$ tal que $D(x, \varepsilon) \subset A$.

6. Determínense, si existen, los límites de las siguientes sucesiones en $\mathbb{R}^2$:

a. $\left((-1)^n, \dfrac{1}{n}\right)$

b. $\left(1, \dfrac{1}{n}\right)$

c. $\left(\left(\dfrac{1}{n}\right)(\cos n\pi), \left(\dfrac{1}{n}\right)\left(\mathrm{sen}\left(n\pi + \dfrac{\pi}{2}\right)\right)\right)$

d. $\left(\dfrac{1}{n}, n^{-n}\right)$

7. Sea U un conjunto abierto en un espacio métrico M. Muéstrese que $U = \mathrm{cl}(U) \setminus \partial U$. ¿Es esto cierto para todo conjunto en M?

8. Sea $S \subset \mathbb{R}$ un conjunto no vacío, acotado inferiormente y cerrado. Muéstrese que $\inf(S) \in S$.

9. Muéstrese que

a. $\mathrm{int}B = B \setminus \partial B$, y

b. $\mathrm{cl}(A) = M \setminus \mathrm{int}(M \setminus A)$.

10. Determínese cuáles de las siguientes proposiciones son verdaderas.

a. $\mathrm{int}(\mathrm{cl}(A)) = \mathrm{int}(A)$.

b. $\mathrm{cl}(A) \cap A = A$.

c. $\mathrm{cl}(\mathrm{int}(A)) = A$.

d. $\partial(\mathrm{cl}(A)) = \partial(A)$.

e. Si A es abierto, entonces $\partial(A) \subset M \setminus A$.

11. Muéstrese que, en un espacio métrico, $x_m \to x$ sii para todo $\varepsilon > 0$ existe N tal que $m \geq N$ implica $d(x_m, x) \leq \varepsilon$ (esto difiere de la proposición **2.7.2** en el hecho de que "$< \varepsilon$" está reemplazado por "$\leq \varepsilon$").

12. Demuéstrense las siguientes propiedades para subconjuntos A y B de un espacio métrico:

a. $\text{int}(\text{int}(A)) = \text{int}(A)$.

b. $\text{int}(A \cup B) \supset \text{int}(A) \cup \text{int}(B)$.

c. $\text{int}(A \cap B) = \text{int}(A) \cap \text{int}(B)$.

13. Muéstrese que $\text{cl}(A) = A \cup \partial(A)$.

14. Demuéstrese lo siguiente para subconjuntos A y B de un espacio métrico M:

a. $\text{cl}(\text{cl}(A)) = \text{cl}(A)$.

b. $\text{cl}(A \cup B) = \text{cl}(A) \cup \text{cl}(B)$.

c. $\text{cl}(A \cap B) \subset \text{cl}(A) \cap \text{cl}(B)$.

15. Demuéstrese lo siguiente para subconjuntos A y B de un espacio métrico M:

a. $\partial(A) = \partial(M \setminus A)$.

b. $\partial(\partial(A)) \subset \partial(A)$.

c. $\partial(A \cup B) \subset \partial(A) \cup \partial(B) \subset \partial(A \cup B) \cup A \cup B$.

d. $\partial(\partial(\partial(A))) = \partial(\partial(A))$.

16. Sean $a_1 = \sqrt{2}$, $a_2 = (\sqrt{2})^{a_1}, \dots, a_{n+1} = (\sqrt{2})^{a_n}$. Muéstrese que $a_n \to 2$ cuando $n \to \infty$. (Se puede usar cualquier propiedad adecuada del cálculo.)

17. Si $\sum x_m$ converge absolutamente en $\mathbb{R}^n$, demuéstrese que $\sum x_m \operatorname{sen} m$ converge.

18. Si $x, y \in M$ y $x \neq y$, entonces demuéstrese que existen sendos conjuntos abiertos U y V tales que $x \in U$, $y \in V$ y $U \cap V = \varnothing$.

19. Definimos un ***punto límite*** de un conjunto A en un espacio métrico M como un punto $x \in M$ tal que $U \cap A \neq \varnothing$ para toda vecindad U de x.

a. ¿Cuál es la diferencia entre los puntos límite y los puntos de acumulación? Proporciónense algunos ejemplos.

b. Si x es un punto límite de A, muéstrese entonces que existe una sucesión $x_n \in A$ tal que $x_n \to x$.

c. Si x es un punto de acumulación de A, muéstrese entonces que x es un punto límite de A. ¿Es cierto el recíproco?

d. Si x es un punto límite de A tal que $x \notin A$, muéstrese entonces que x es un punto de acumulación.

e. Demuéstrese que un conjunto es cerrado sii contiene todos sus puntos límite.

20. Dado un conjunto A en un espacio métrico M y $x \in M$, sea

$$d(x,A) = \inf\{d(x,y) \mid y \in A\},$$

y dado $\varepsilon > 0$, sea $D(A,\varepsilon) = \{x \mid d(x,A) < \varepsilon\}$.

a. Muéstrese que $D(A,\varepsilon)$ es abierto.

b. Sean $A \subset M$ y $N_\varepsilon = \{x \in M \mid d(x,A) \leq \varepsilon\}$, donde $\varepsilon > 0$. Muéstrese que N_ε es cerrado y que A es cerrado sii $A = \bigcap \{N_\varepsilon \mid \varepsilon > 0\}$.

21. Demuéstrese que una sucesión x_k en un espacio normado es una sucesión de Cauchy sii para toda vecindad U de 0 existe N tal que $k, l \geq N$ implica que $x_k - x_l \in U$.

22. Demuéstrese la proposición **2.3.2**. (Sugerencia: utilícese el ejercicio 12 de la Introducción.)

23. Demuéstrese que el interior de un conjunto $A \subset M$ es la unión de todos los subconjuntos abiertos de A. Dedúzcase que A es abierto sii $A = \text{int}(A)$. Proporciónese además una demostración directa de la última afirmación, a partir de las definiciones.

24. Identifíquese $\mathbb{R}^{n+m}$ con $\mathbb{R}^n \times \mathbb{R}^m$. Muéstrese que $A \subset \mathbb{R}^{n+m}$ es abierto sii para cada $(x, y) \in A$, con $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, existen sendos conjuntos abiertos $U \subset \mathbb{R}^n$ y $V \subset \mathbb{R}^m$ tales que $x \in U$, $y \in V$ y $U \times V \subset A$. Dedúzcase que el producto de conjuntos abiertos es abierto.

25. Demuéstrese que un conjunto $A \subset M$ es abierto sii se puede escribir como la unión de alguna familia de discos de radio ε.

26. Definamos la sucesión de números a_n como

$$a_0 = 1,\ a_1 = 1 + \frac{1}{1+a_0},\ \ldots,\ a_n = 1 + \frac{1}{1+a_{n-1}}.$$

Muéstrese que a_n es una sucesión convergente y determínese su límite.

27. Supongamos que $a_n \geq 0$ y que $a_n \to 0$ cuando $n \to \infty$. Dado $\varepsilon > 0$, muéstrese que existe una subsucesión b_n de a_n tal que $\sum_{n=1}^{\infty} b_n < \varepsilon$.

28. Encuéntrense ejemplos de:

a. Un conjunto infinito en $\mathbb{R}$ sin puntos de acumulación.

b. Un subconjunto no vacío de $\mathbb{R}$ contenido en su conjunto de puntos de acumulación

c. Un subconjunto de $\mathbb{R}$ con infinitos puntos de acumulación, pero que no contenga ninguno de ellos

d. Un conjunto A tal que $\partial(A) = \text{cl}(A)$.

29. Sean $A, B \subset \mathbb{R}^n$ y x un punto de acumulación de $A \cup B$. ¿Debe x ser un punto de acumulación de A o de B?

30. Muéstrese que cualquier conjunto abierto en $\mathbb{R}$ es una unión de intervalos abiertos disjuntos. ¿Se verificará lo mismo en $\mathbb{R}^n$ para $n > 1$, si definimos un intervalo abierto como el producto cartesiano de n intervalos abiertos, $]a_1, b_1[\times \cdots \times]a_n, b_n[$?

31. Sea A' el conjunto de puntos de acumulación de un conjunto A. Demuéstrese que A' es cerrado. ¿Es cierto que $(A')' = A'$ para todo A?

32. Sean $A \subset \mathbb{R}^n$ un conjunto cerrado y $x_n \in A$ una sucesión de Cauchy. Demuéstrese que x_n converge a un punto en A.

33. Sea s_n una sucesión acotada de números reales. Supongamos que $2s_n \leq s_{n-1} + s_{n+1}$. Muéstrese que $\lim_{n\to\infty} (s_{n+1} - s_n) = 0$.

34. Sean $x_n \in \mathbb{R}^k$ y $d(x_{n+1}, x_n) \leq rd(x_n, x_{n-1})$, donde $0 \leq r < 1$. Muéstrese que x_n converge.

35. Muéstrese que cualquier familia de conjuntos abiertos no vacíos disjuntos de números reales es numerable.

36. Sean $A, B \subset \mathbb{R}^n$ conjuntos cerrados. ¿Debe $A + B = \{x + y \mid x \in A \text{ e } y \in B\}$ ser cerrado?

37. Sea $A \subset M$ un espacio métrico. Demuéstrese que

$$\partial(A) = [A \cap \text{cl}(M \setminus A)] \cup [\text{cl}(A) \setminus A].$$

38. Sea $x_k \in \mathbb{R}^n$ tal que $\|x_k - x_l\| \leq 1/k + 1/l$. Demuéstrese que x_k converge.

39. Sea $S \subset \mathbb{R}$ un conjunto acotado superior e inferiormente. Demuéstrese que $\sup(S) - \inf(S) = \sup\{x - y \mid x \in S \text{ e } y \in S\}$.

40. Supongamos que para todo n se tiene $a_n \leq b_n$, $a_n \leq a_{n+1}$ y $b_{n+1} \leq b_n$, donde a_n y b_n son sucesiones de números reales. Demuéstrese que a_n converge.

41. Sea A_n una colección de subconjuntos de un espacio métrico M, tales que $A_{n+1} \subset A_n$ y que $A_n \neq \varnothing$; supongamos que $\bigcap_{n=1}^{\infty} A_n \neq \varnothing$. Sea $x \in \bigcap_{n=1}^{\infty} \text{cl}(A_n)$; muéstrese que x es un punto de acumulación de A_1.

42. Sean $A \subset \mathbb{R}^n$ y $x \in \mathbb{R}^n$. Definamos $d(x, A) = \inf\{d(x, y) \mid y \in A\}$. ¿Existirá $z \in A$ tal que $d(x, A = d(x, z)$?

43. Sean $x_1 = \sqrt{3}, \ldots, x_n = \sqrt{3 + x_{n-1}}$. Calcúlese $\lim_{n\to\infty} x_n$.

44. Se dice que un conjunto $A \subset \mathbb{R}^n$ es ***denso*** en $B \subset \mathbb{R}^n$ si $B \subset \mathrm{cl}(A)$. Si A es denso en $\mathbb{R}^n$ y U es abierto, demuéstrese que $A \cap U$ es denso en U. ¿Se cumple también si U no es abierto?

45. Muéstrese que $x^{\log x} = o(e^x)$ cuando $x \to \infty$ (véase el ejemplo resuelto **2.5**).

46. **a.** Si $f = o(g)$ y $g(x) \to 0$ cuando $x \to \infty$, muéstrese entonces que $e^f = o(e^g)$.

b. Muéstrese que $\lim_{n\to\infty} (x \log x)/e^x = 0$, probando que $x = o(e^{x/2})$ y que $\log x = o(e^{x/2})$.

47. Muéstrese que $\gamma = \lim_{n\to\infty} (1 + \frac{1}{2} + \cdots + \frac{1}{2} - \log n)$ existe usando la demostración del criterio de la integral (γ es la ***constante de Euler***).

48. Demuéstrense las siguientes generalizaciones de los criterios del cociente y de la raíz:

a. Si $a_n > 0$ y $\limsup_{n\to\infty} a_{n+1}/a_n < 1$, entonces converge, y si $\liminf_{n\to\infty} a_{n+1}/a_n > 1$, entonces $\sum a_n$ diverge.

b. Si $a_n \geq 0$ y $\limsup_{n\to\infty} \sqrt[n]{a_n} < 1$ (respectivamente, > 1), entonces $\sum a_n$ converge (respectivamente, diverge).

c. En el criterio de comparación en el límite, ¿podemos reemplazar los límites por límites superiores?

49. Demuéstrese el ***criterio de Raabe***: si $a_n > 0$ y $a_{n+1}/a_n \leq 1 - A/n$ para alguna constante $A > 1$ y para n suficientemente grande, entonces $\sum a_n$ converge. De forma análoga, muéstrese que si $a_{n+1}/a_n \geq 1 - (1/n)$, entonces $\sum a_n$ diverge.

Utilícese el criterio de Raabe para demostrar la convergencia de la ***serie hipergeométrica***, cuyo término general es

$$a_n = \frac{\alpha(\alpha+1)\cdots(\alpha+n-1)\beta(\beta+1)\cdots(\beta+n-1)}{1\cdot 2\cdots n\cdot\gamma(\gamma+1)\cdots(\gamma+n-1)}$$

donde α, β y γ son enteros no negativos tales que $\gamma > \alpha + \beta$. Muéstrese que la serie diverge si $\gamma < \alpha + \beta$.

50. Muéstrese que para x suficientemente grande, $f(x) = (x\cos^2 x + \operatorname{sen}^2 x)e^{x^2}$ es monótona y tiende a $+\infty$, pero que ni el cociente $f(x)/(x^{1/2}e^{x^2})$ ni su recíproco están acotados.

51. **a.** Si $u_n > 0$, $n = 1, 2, \ldots$, muéstrese que

$$\liminf \frac{u_{n+1}}{u_n} \leq \liminf \sqrt[n]{u_n} \leq \limsup \sqrt[n]{u_n} \leq \limsup \frac{u_{n+1}}{u_n}.$$

b. Dedúzcase que si $\lim(u_{n+1}/u_n) = A$, entonces $\limsup \sqrt[n]{u_n} = A$.

c. Muéstrese que el recíproco del apartado **b** es falso empleando la sucesión $u_{2n} = u_{2n+1} = 2^{-n}$.

d. Calcúlese $\limsup \sqrt[n]{n!}\,/\,n$.

52. Verifíquese la convergencia o divergencia de las siguientes series.

a. $\displaystyle\sum_{k=0}^{\infty} \frac{e^{-k}}{\sqrt{k+1}}$

b. $\displaystyle\sum_{k=0}^{\infty} \frac{k}{k^2+1}$

c. $\displaystyle\sum_{n=0}^{\infty} \frac{\sqrt{n+1}}{n^2-3n+1}$

d. $\displaystyle\sum_{k=1}^{\infty} \frac{\log(k+1)-\log k}{\tan^{-1}(2/k)}$

e. $\displaystyle\sum_{n=1}^{\infty} \operatorname{sen}(n^{-\alpha})$, α real, > 0

f. $\displaystyle\sum_{n=1}^{\infty} \frac{n^3}{3^n}$

53. Dado un conjunto A en un espacio métrico, ¿cuál es el número máximo de subconjuntos que se pueden obtener mediante la aplicación sucesiva de las operaciones de clausura, interior y complemento a A (en cualquier orden)? Proporciónese un ejemplo de un conjunto que alcance tal máximo.

Capítulo 3
Conjuntos compactos y conexos

En este capítulo estudiaremos dos de los tipos de conjuntos más importantes y útiles en los espacios métricos y en particular en $\mathbb{R}^n$. Intuitivamente, entendemos que un conjunto en $\mathbb{R}^n$ es compacto cuando es cerrado y está contenido en una región acotada, y que es conexo, cuando el conjunto es "de una pieza". La figura 3-1 muestra algunos ejemplos. Como de costumbre, es necesario transformar estas ideas en definiciones rigurosas. En cada caso, la definición técnica más útil parece un poco apartada de nuestra intuición, pero al final veremos que cuadra con ella. En el capítulo 4 revelaremos lo fructífero de estos conceptos, cuando los apliquemos al estudio de las funciones continuas.

§3.1 Compacidad

En esta sección daremos la definición general y las propiedades de los conjuntos compactos en espacios métricos. Un criterio para reconocer los conjuntos compactos, el ***teorema de Heine-Borel***, establece que un conjunto en $\mathbb{R}^n$ es compacto si es cerrado y acotado. Analizaremos este resultado, particular del espacio métrico $\mathbb{R}^n$, en §3.2.

Recordemos de nuestro estudio de la completitud de $\mathbb{R}^n$ en el capítulo 1 que toda sucesión acotada tiene una subsucesión convergente. Podemos expresar este hecho de otra forma: si $A \subset \mathbb{R}^n$ es un conjunto cerrado y acotado, entonces toda sucesión en A tiene una subsucesión que converge a un punto de A. Desde un punto de vista histórico, este resultado fue identificado como una propiedad importante de los conjuntos, y por ello fue elevado al rango de definición. Dicha propiedad desempeña un papel crucial en muchos teoremas básicos, como la existencia de máximos y mínimos de funciones continuas en intervalos cerrados, según veremos en el capítulo 4.

3.1.1 Definición *Sea M un espacio métrico. Un subconjunto $A \subset M$ es **secuencialmente compacto** si toda sucesión en A tiene una subsucesión que converge a un punto de A.*

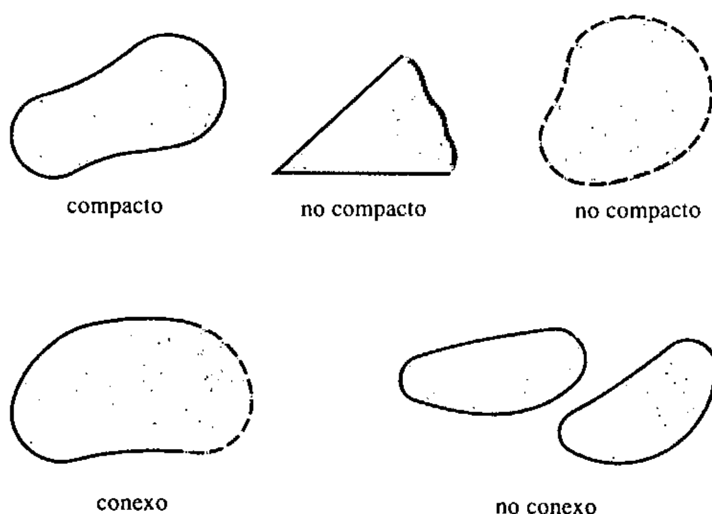


FIGURA 3-1 Conjuntos compactos y conexos en $\mathbb{R}^2$

Esta propiedad es equivalente a otra, llamada *compacidad*, que desarrollaremos a continuación. Dicha propiedad es menos evidente y su equivalencia con la compacidad secuencial está lejos de ser inmediata, al menos a primera vista.

Vamos a introducir cierta terminología que necesitaremos para nuestra definición formal. Sea M un espacio métrico y $A \subset M$ un subconjunto. Un ***recubrimiento*** de A es una colección $\{U_i\}$ de conjuntos cuya unión contiene a A; es un ***recubrimiento abierto*** si cada U_i es abierto. Un ***subrecubrimiento*** de un recubrimiento dado es una subcolección de $\{U_i\}$ cuya unión también contiene a A o, como suele decirse, ***recubre*** A; es un ***subrecubrimiento finito*** si la subcolección sólo contiene un número finito de conjuntos.

Los recubrimientos abiertos no son necesariamente colecciones numerables de conjuntos abiertos. Por ejemplo, el conjunto no numerable de discos $\{D((x, 0), 1) \mid x \in \mathbb{R}\}$ en $\mathbb{R}^2$ recubre el eje real, y la subcolección de todos los discos $D((n, 0), 1)$ con centro en los puntos enteros de la recta real forma un subrecubrimiento numerable. Nótese que el conjunto de discos $D((2n, 0), 1)$ con centro en los puntos enteros pares de la recta real no forma un subrecubrimiento (¿por qué?).

3.1.2 Definición *Un subconjunto A de un espacio métrico M es **compacto** si todo recubrimiento abierto de A contiene un subrecubrimiento finito.*

Y aquí tenemos el primer resultado importante, que relaciona la compacidad con la compacidad secuencial.

3.1.3 Teorema de Bolzano-Weierstrass *Un subconjunto de un espacio métrico es compacto sii es secuencialmente compacto.*

Unas cuantas observaciones elementales nos ayudarán a dar una idea del concepto de compacidad y de este teorema. En primer lugar, *un conjunto secuencialmente compacto debe ser cerrado*. De hecho, si $x_n \in A$ converge a $x \in M$, entonces, por hipótesis, existe una subsucesión convergente a un punto $x_0 \in A$; por la unicidad de los límites, $x = x_0$, de modo que A es cerrado. En segundo lugar, un conjunto secuencialmente compacto A *debe estar acotado*; de lo contrario, existirían un punto $x_0 \in A$ y una sucesión $x_n \in A$ tales que $d(x_n, x_0) \geq n$, con lo que x_n no podría contener una subsucesión convergente. Para probar directamente que un conjunto compacto está acotado, se puede usar el hecho de que para cualquier $x_0 \in A$, las bolas abiertas $D(x_0, n)$, $n = 1, 2, \ldots$ recubren A, por lo que existirá un subrecubrimiento finito.

Obsérvese que en las definiciones se puede tomar $A = M$, en cuyo caso se habla de un ***espacio métrico compacto***. En su momento elaboraremos algunos ejemplos de espacios compactos.

Otra caracterización de la compacidad se relaciona con la completitud y resulta una herramienta técnica útil en la demostración del teorema de Bolzano-Weierstrass.

3.1.4 Definición *Un conjunto $A \subset M$ es **totalmente acotado** si para todo $\varepsilon > 0$ existe un conjunto finito $\{x_1, \ldots, x_N\}$ en M tal que $A \subset \bigcup_{i=1}^{N} D(x_i, \varepsilon)$.*

3.1.5 Teorema *Un espacio métrico es compacto sii es completo y totalmente acotado.*

Sea $A \subset M$ y supongamos que M es completo. Si aplicamos este teorema al espacio métrico A, podemos concluir que *A es compacto sii es cerrado y totalmente acotado*.

En el teorema **3.1.5**, algunas cosas son evidentes y otras no. En primer lugar, obsérvese que $D(x_i, \varepsilon) \subset D(x_1, \varepsilon + d(x_i, x_1))$, de modo que si

$$R = \varepsilon + \text{máx}\{d(x_2, x_1), \ldots, d(x_N, x_1)\},$$

entonces $A \subset D(x_1, R)$ y, por lo tanto, *un conjunto totalmente acotado está acotado*. Esto es consistente con nuestra observación anterior de que los conjuntos compactos están acotados.

En este momento no disponemos de métodos eficaces que nos digan si un conjunto dado es compacto o no, pero lo remediaremos en la siguiente sección.

3.1.6 Ejemplo La recta real completa $\mathbb{R}$ no es compacta, ya que no está acotada. Otra razón es que

$$\{D(n, 1) =]n - 1, n + 1[\mid n = 0, \pm 1, \pm 2, \ldots\}$$

es un recubrimiento abierto de $\mathbb{R}$ que no contiene un subrecubrimiento finito (¿por qué?). ♦

3.1.7 Ejemplo *Sea $A =]0, 1]$. Determínese un recubrimiento abierto que no contenga un subrecubrimiento finito.*

Solución Considérese el recubrimiento abierto $\{]1/n, 2[\mid n = 1, 2, 3, \ldots\}$. (¿Por qué la unión contiene todo A?) Está claro que no puede contener un subrecubrimiento finito. Esta vez, la compacidad falla porque A no es cerrado: "falta" en A el punto 0. Esta colección no es un recubrimiento de $[0, 1]$; de hecho, cualquier recubrimiento abierto de $[0, 1]$ debe contener un subrecubrimiento finito, pues, como demostraremos en la siguiente sección, $[0, 1]$ es compacto. ♦

3.1.8 Ejemplo *Encuéntrese un ejemplo de un conjunto acotado y cerrado que no sea compacto.*

Solución Sea M cualquier conjunto infinito con la métrica discreta: $d(x, y) = 0$ si $x = y$ y $d(x, y) = 1$ si $x \neq y$. Claramente $M \subset D(x_0, 2)$ para todo $x_0 \in M$, de modo que M está acotado. Como ya es el espacio total, es cerrado. Sin embargo, no es compacto: $\{D(x, 1/2) \mid x \in M\}$ es un recubrimiento abierto que no contiene un subrecubrimiento finito. ♦

3.1.9 Ejemplo *Una colección de conjuntos cerrados $\{K_\alpha\}$ en un espacio métrico M tiene la* ***propiedad de intersección finita*** *para A si la intersección de cualquier número finito de K_α con A es no vacía. Muéstrese que $A \subset M$ es compacto sii toda colección de conjuntos cerrados con la propiedad de intersección finita para A tiene intersección no vacía con A.*

Solución En primer lugar supondremos que A es compacto. Sea $\{F_i\}$ una colección de conjuntos cerrados y sea $U_i = M \backslash F_i$, de modo que U_i es abierto. Supongamos que $A \cap (\cap_{i=1}^{\infty} F_i) = \varnothing$. Tomando complementarios, esto quiere decir que los conjuntos U_i recubren A. Puesto que el recubrimiento es abierto, existe un subrecubrimiento finito, digamos, $A \subset U_1 \cup \cdots \cup U_N$. Entonces $A \cap (F_1 \cap \cdots \cap F_N) = \varnothing$ y, por lo tanto, $\{F_i\}$ no tiene la propiedad de intersección finita. Así, si $\{F_i\}$ es una colección de conjuntos cerrados con la propiedad de intersección finita, entonces $A \cap \{F_i\} \neq \varnothing$.

Recíprocamente, sea $\{U_i\}$ un recubrimiento abierto de A y sea $F_i = M \backslash U_i$. Entonces $A \cap (\cap_{i=1}^{\infty} F_i) = \varnothing$ y, por hipótesis, $\{F_i\}$ no puede tener la propiedad de intersección finita para A. Así, $A \cap (F_i \cap \cdots \cap F_N) = \varnothing$ para unos ciertos elementos $(F_1, \ldots, F_N)$ de la colección. Por lo tanto, $U_1, \ldots, U_N$ es el subrecubrimiento finito necesario y por ello A es compacto. ♦

Ejercicios de §3.1

1. Muéstrese que $A \subset M$ es secuencialmente compacto sii todo subconjunto infinito de A tiene un punto de acumulación en A.

2. Demuéstrese que $\{(x, y) \in \mathbb{R}^2 \mid 0 \leq x < 1, 0 \leq y \leq 1\}$ no es compacto.

3. Sea M un espacio completo y $A \subset M$ totalmente acotado. Muéstrese que cl(A) es compacto.

4. Sea $x_k \to x$ una sucesión convergente en un espacio métrico y sea $A = \{x_1, x_2, \ldots\} \cup \{x\}$.

- **a.** Muéstrese que A es compacto.
- **b.** Verifíquese que todo recubrimiento abierto de A contiene un subrecubrimiento finito.

5. Sea M un conjunto con la métrica discreta. Muéstrese que ningún subconjunto infinito de M es compacto. ¿Por qué no contradice esto la afirmación del ejercicio 4?

§3.2 Teorema de Heine-Borel

En el espacio euclídeo, podemos decir fácilmente si un conjunto es compacto mediante el siguiente teorema:

3.2.1 Teorema de Heine-Borel *Un conjunto $A \subset \mathbb{R}^n$ es compacto sii es cerrado y acotado.*

En §**3.1** ya mencionamos la mitad de este teorema. De hecho, un conjunto compacto es cerrado y acotado en *cualquier* espacio métrico. El recíproco debe ser especial, en vista del ejemplo **3.1.8**. De hecho, ni siquiera es evidente que el intervalo cerrado $[0, 1]$ en $\mathbb{R}$ sea compacto. En realidad sí lo es, y una de las demostraciones del teorema de Heine-Borel comienza analizando este caso.

3.2.2 Ejemplo *Determínese cuáles de los siguientes conjuntos son compactos:*

- **a.** $\{x \in \mathbb{R} \mid x \geq 0\} \subset \mathbb{R}$
- **b.** $[0, 1] \cup [2, 3] \subset \mathbb{R}$
- **c.** $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\} \subset \mathbb{R}^2$

Solución

a. No es compacto, porque no está acotado.

b. Compacto, porque es cerrado y acotado.

c. No es compacto, porque no es cerrado. ♦

3.2.3 Ejemplo *Sea x_k una sucesión de puntos en $\mathbb{R}^n$ tales que $\|x_k\| \leq 3$ para todo k. Muéstrese que x_k tiene una subsucesión convergente.*

Solución El conjunto $A = \{x \in \mathbb{R}^n \mid \|x\| \leq 3\}$ es cerrado y acotado, por lo que es compacto. Como $x_k \in A$, podemos aplicar el teorema de Bolzano-Weierstrass para obtener la conclusión. ♦

3.2.4 Ejemplo *En la definición de un conjunto compacto, ¿podemos reemplazar la palabra "todo" por "algún"?*

Solución No: sea $A = \mathbb{R}$ y consideremos el recubrimiento abierto formado solamente por el conjunto abierto $\mathbb{R}$. Éste contiene un subrecubrimiento finito (él mismo), pero al no ser un conjunto acotado, $\mathbb{R}$ no es compacto. ♦

3.2.5 Ejemplo *Sea $A = \{0\} \cup \{1, 1/2, \ldots, 1/n, \ldots\}$. Muéstrese directamente que A satisface la definición de compacidad.*

Solución Sea $\{U_i\}$ un recubrimiento abierto arbitrario de A. Debemos mostrar que contiene un subrecubrimiento finito. El punto 0 estará en uno de los conjuntos abiertos; cambiando los índices podemos suponer que $0 \in U_1$. Como U_1 es abierto y $1/n \to 0$, existirá N tal que $1/N, 1/(N+1), \ldots$ están en U_1. Cambiemos los nombres de nuevo en caso necesario para poder suponer que $1 \in U_2, \ldots, 1/(1-N) \in U_N$. Entonces, $U_1, \ldots, U_N$ es un subrecubrimiento finito, pues es una subcolección finita de $\{U_i\}$ que contiene todos los puntos de A. Obsérvese que si A fuera el conjunto $\{1, 1/2, \ldots\}$, no funcionaría el argumento; de hecho, este conjunto no es cerrado, por lo cual no es compacto. ♦

Ejercicios de §3.2

1. ¿Cuáles de los siguientes conjuntos son compactos?

a. $\{x \in \mathbb{R} \mid 0 \leq x \leq 1 \text{ y } x \text{ es irracional}\}$

b. $\{(x,y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1\}$

c. $\{(x,y) \in \mathbb{R}^2 \mid xy \geq 1\} \cap \{(x,y) \mid x^2 + y^2 < 5\}$

2. Sea $r_1, r_2, r_3, \ldots$ una enumeración de los números racionales en $[0, 1]$. Muéstrese que existe una subsucesión convergente.

3. Sea $M = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ con la métrica usual. Muéstrese que $A \subset M$ es compacto sii A es cerrado.

4. Sea A un conjunto acotado en $\mathbb{R}^n$. Demuéstrese que $\mathrm{cl}(A)$ es compacto.

5. Sea A un conjunto infinito en $\mathbb{R}$ con un único punto de acumulación en A. ¿Debe A ser compacto?

§3.3 Propiedad de los conjuntos encajados

El siguiente teorema es una consecuencia importante del teorema de Bolzano-Weierstrass.

3.3.1 Propiedad de los conjuntos encajados *Sea F_k una sucesión de conjuntos compactos no vacíos en un espacio métrico M tales que $F_{k+1} \subset F_k$ para todo $k =$* 1, 2, ... Entonces existe al menos un punto en $\bigcap_{k=1}^{\infty} F_k$.

Intuitivamente, los conjuntos F_k son no vacíos y decrecientes, de modo que parece razonable que exista un punto en todos ellos. Sin embargo, si los conjuntos F_k no fueran compactos, entonces la intersección podría ser vacía (véase el ejemplo **3.3.4**). Así pues, la demostración requiere un poco más de cuidado.

Para demostrar la propiedad de los conjuntos encajados mediante el teorema de Bolzano-Weierstrass, elegimos $x_k \in F_k$ para cada k. La sucesión x_k tiene una subsucesión convergente, pues está contenida en el conjunto compacto F_1. El punto límite está en todos los conjuntos F_k porque éstos son cerrados (véase la figura 3.3-1). Al final del capítulo daremos una demostración alternativa.

La propiedad de los conjuntos encajados puede plantearse en términos de "conjuntos crecientes" de la siguiente forma. Sea $U_k = M \backslash F_k$, de modo que los conjuntos U_k son abiertos y $U_{k+1} \supset U_k$. Entonces $\bigcup_{k=1}^{\infty} U_k \neq M$ es equivalente a $\bigcap_{k=1}^{\infty} F_k \neq \varnothing$. Así, *si M es un espacio métrico, los conjuntos abiertos U_k son crecientes (es decir, $U_{k+1} \supset U_k$) y sus complementarios son compactos, por lo que la unión de los conjuntos U_k no es todo M.*

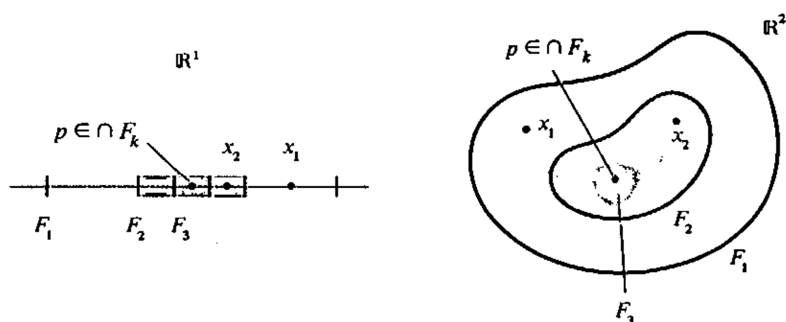


FIGURA 3.3-1 Propiedad de los conjuntos encajados

3.3.2 Ejemplo Sea M la esfera unidad en $\mathbb{R}^3$, $M = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}$, con la métrica usual. Sea U_i la parte de M que está estrictamente debajo de la latitud $90° - 10/i$, $i = 1, 2, 3, \ldots$, como muestra la figura 3.3-2.

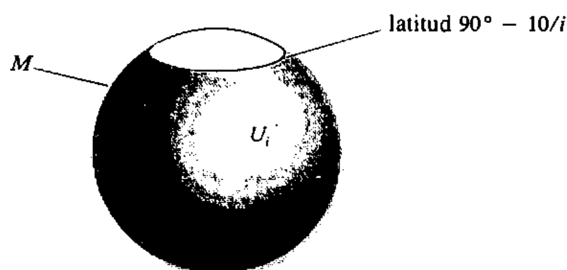


FIGURA 3.3-2 Una sucesión creciente de conjuntos abiertos sobre la esfera unidad

El espacio métrico M es compacto (¿por qué?) y, consistente con las observaciones anteriores, la unión de los U_i no es todo M, ya que no incluye el polo norte. ♦

3.3.3 Ejemplo *Verifíquese la propiedad de los conjuntos encajados para* $F_k = [0, 1/k] \subset \mathbb{R}$.

Solución Cada F_k es compacto y $F_{k+1} \subset F_k$. La intersección es $\{0\}$, que no es vacía. ♦

3.3.4 Ejemplo *¿Es cierta la propiedad de los conjuntos encajados si reemplazamos "compactos no vacíos" por "abiertos no vacíos" o "cerrados no vacíos"?*

Solución No. Sean $F_k =]k, \infty[$ o $[k, \infty[$. ♦

3.3.5 Ejemplo Una familia más exótica de conjuntos compactos decrecientes F_n para la que el conjunto $\bigcap_{n=1}^{\infty} F_n$ es un poco complicado se obtiene al eliminar sucesivamente triángulos de un triángulo dado en el plano, como muestra la figura 3.3-3. ♦

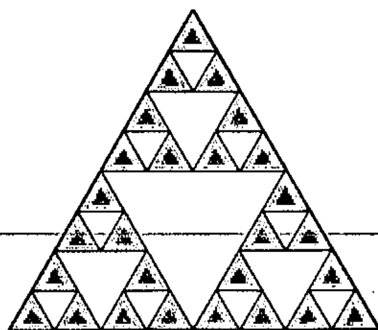


FIGURA 3.3-3 Cesta de Sierpinski

Ejercicios de §3.3

1. Verifíquese la propiedad de los conjuntos encajados para $F_k = \{x \in \mathbb{R} \mid x \geq 0, 2 \leq x^2 \leq 2 + 1/k\}$.

2. ¿Es cierta la propiedad de los conjuntos encajados si reemplazamos "compactos no vacíos" por "abiertos acotados no vacíos"?

3. Sea $x_k \to x$ una sucesión convergente en un espacio métrico. Verifíquese la validez de la propiedad de los conjuntos encajados para $F_k = \{x_l \mid l \leq k\} \cup \{x\}$. ¿Qué ocurre si $F_k = \{x_l \mid l \geq k\}$?

4. Sean $x_k \to x$ una sucesión convergente en un espacio métrico y $\mathcal{A}$ una familia de conjuntos cerrados con la propiedad de que para cada $A \in \mathcal{A}$, existe N tal que $k \geq N$ implica $x_k \in A$. Demuéstrese que $x \in \bigcap \mathcal{A}$.

§3.4 Conjuntos conexos por arcos

El segundo tema importante que analizaremos en este capítulo es la conexión. Sabemos intuitivamente a qué conjuntos queremos llamar "conexos", sin embargo, nuestra intuición puede fallar al juzgar conjuntos más complicados. Por ejemplo, ¿cómo decidimos si el conjunto $\{(x, \text{sen}(1/x)) \mid x > 0\} \cup \{(0, y) \mid y \in [-1, 1]\} \subset \mathbb{R}^2$ es conexo (véase la figura 3.4-1)? Por lo tanto, buscaremos una definición matemática firme en la cual basarnos.

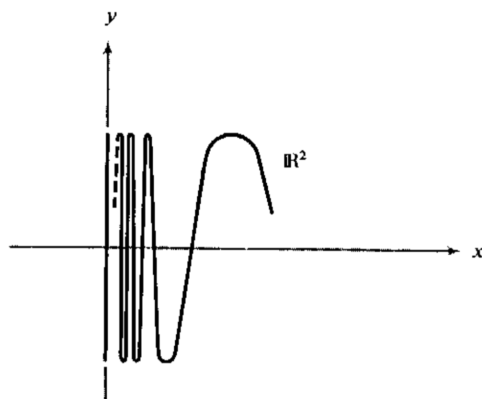


FIGURA 3.4-1 ¿Es conexo?

Existen, de hecho, dos nociones de conexión diferentes (aunque íntimamente relacionadas), la más intuitiva y aplicable de las cuales es la conexión por arcos, por lo que comenzamos por ella. Primero debemos definir lo que significa que una curva (o arco) una dos puntos.

3.4.1 Definición *Decimos que una aplicación $\varphi : [a, b] \to M$ de un intervalo $[a, b]$ en un espacio métrico M es **continua** si $(t_k \to t)$ implica $(\varphi(t_k) \to \varphi(t))$ para toda sucesión t_k en $[a, b]$ convergente a algún $t \in [a, b]$. (El lector recordará de su curso de cálculo que, intuitivamente, una función continua no tiene "interrupciones" o "saltos" en su gráfica.) Un **arco continuo** que une dos puntos, x, y, en un espacio métrico M es una aplicación $\varphi : [a, b] \to M$ tal que $\varphi(a) = x$, $\varphi(b) = y$, y φ es continua. Aquí, x puede, o no, ser igual a y, y $b \geq a$. Un arco φ **está contenido en un conjunto** A si $\varphi(t) \in A$ para todo $t \in [a, b]$. Véase la figura 3.4-2.*

Decimos que un conjunto es **conexo por arcos** si cualesquiera dos puntos del conjunto se pueden unir mediante un arco continuo contenido en el conjunto.

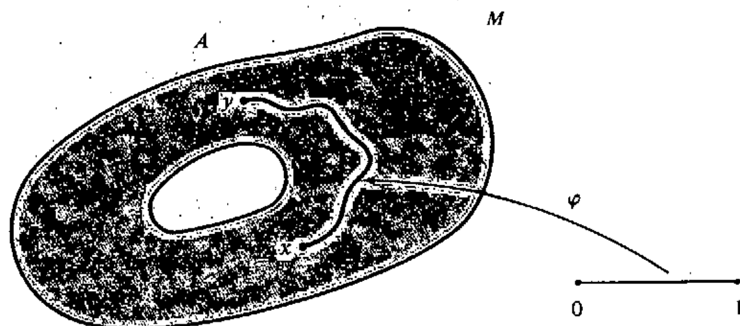


FIGURA 3.4-2 Una curva que une x e y en A

Por ejemplo, es evidente que la región A de la figura 3.4-2 es conexa por arcos. Otro conjunto conexo por arcos es el intervalo [0, 1]. Para demostrarlo, sean $x, y \in [0, 1]$ y definamos $\varphi : [0, 1] \to \mathbb{R}$ como $\varphi(t) = (y - x)t + x$. Éste es un arco continuo que une x e y, contenido en [0, 1].

Si usamos la definición anterior de conexión por arcos, un poco de reflexión convencerá al lector de que el conjunto de la figura 3.4-1 no es conexo por arcos, aunque este hecho no es obvio. Con frecuencia será fácil determinar si un conjunto es conexo por arcos viendo si dos puntos cualesquiera se pueden unir mediante una curva continua contenida en el conjunto, lo que, por lo general, es claro geométricamente. La segunda idea de conexión es más difícil de verificar directamente, pero nos será de mucha utilidad; ésta aparece definida en **§3.5.**

3.4.2 Ejemplo *¿Cuáles de los siguientes conjuntos son conexos por arcos?*

a. $[0,3]$

b. $[1,2] \cup [3,4]$

c. $\{(x,y) \in \mathbb{R}^2 \mid 0 < x \leq 1\}$

d. $\{(x,y) \in \mathbb{R}^2 \mid 0 < x^2 + y^2 \leq 1\}$

Solución Solamente **b** no es conexo por arcos, como resulta evidente de analizar la figura 3.4-3. ♦

3.4.3 Ejemplo *¿Debe un conjunto conexo por arcos ser abierto o cerrado?*

Solución No; por ejemplo, [0, 1],]0, 1[y [0, 1[son todos conexos por arcos. ♦

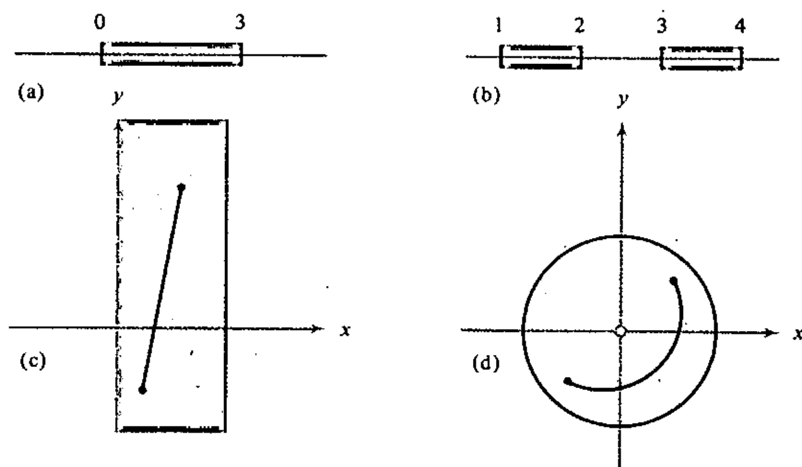


FIGURA 3.4-3 Conjuntos del ejemplo **3.4.2**

3.4.4 Ejemplo *Sea* $\varphi : [0, 1] \to \mathbb{R}^2$ *un arco continuo y* $C = \varphi([0, 1])$. *Muéstrese que* C *es conexo por arcos.*

Solución Esto resulta intuitivamente claro, ya que podemos usar el propio arco φ para unir dos puntos en C. Concretamente, si $x = \varphi(a)$, $y = \varphi(b)$, donde $0 \leq a \leq b \leq 1$, sea $c : [a, b] \to \mathbb{R}^2$ dada por $c(t) = \varphi(t)$. Entonces c es un arco que une x e y, contenido en C. ♦

Ejercicios de §3.4

1. Determínense cuáles de los siguientes conjuntos son conexos por arcos:

a. $\{x \in [0, 1] \mid x \text{ es racional}\}$

b. $\{(x,y) \in \mathbb{R}^2 \mid xy \geq 1 \;\text{ y }\; x > 1\} \cup \{(x,y) \in \mathbb{R}^2 \mid xy \leq 1 \;\text{ y }\; x \leq 1\}$

c. $\{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 \leq z\} \cup \{(x,y,z) \mid x^2 + y^2 + z^2 > 3\}$

d. $\{(x,y) \in \mathbb{R}^2 \mid 0 \leq x < 1\} \cup \{(x,0) \mid 1 < x < 2\}$

2. Sea $A \subset \mathbb{R}$ conexo por arcos. Proporciónese un argumento plausible para mostrar que A debe ser un intervalo (cerrado, abierto o semiabierto). ¿Es todo igual de sencillo en $\mathbb{R}^2$?

3. Sean $\varphi : [a, b] \to \mathbb{R}^3$ un arco continuo, $a < c < d < b$ y $C = \{\varphi(t) \mid c \leq t \leq d\}$. ¿Debe $\varphi^{-1}(C)$ ser conexo por arcos?

§3.5 Conjuntos conexos

Hay otra forma de decir que un conjunto tiene más de una pieza, más sofisticada pero también más poderosa que la conexión por arcos.

3.5.1 Definición *Sea A un subconjunto de un espacio métrico M. Dos conjuntos abiertos U,V* ***separan*** *A si satisfacen estas condiciones:*

a. $U \cap V \cap A = \varnothing$.

b. $A \cap U \neq \varnothing$.

c. $A \cap V \neq \varnothing$.

d. $A \subset U \cup V$.

Véase la figura 3.5-1. Decimos que A es ***disconexo*** *si existen tales conjuntos; en caso contrario, decimos que A es* ***conexo****.*

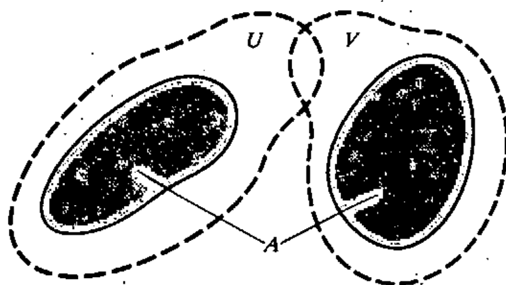


FIGURA 3.5-1 *A* no es ni conexo ni conexo por arcos

Se puede demostrar que el conjunto de la figura 3.4-1 es conexo pero no conexo por arcos; así, estos dos conceptos son diferentes; sin embargo, existe una relación válida entre los dos, que presentamos en el siguiente teorema.

3.5.2 Teorema *Los conjuntos conexos por arcos son conexos.*

El uso de este teorema es tal vez la forma más sencilla de identificar un conjunto conexo. Este teorema es razonable, intuitivamente; de hecho, el teorema recíproco (falso) también es "razonable". Éste es, pues, un ejemplo de dos conceptos que están muy relacionados y que intuitivamente son casi idénticos, pero cuya verdadera relación debe discernirse con más cuidado.

Si un conjunto no es conexo (y, por lo tanto, no es conexo por arcos), podemos separarlo en dos partes, o componentes. Más concretamente, una ***componente conexa*** de un conjunto A es un subconjunto conexo $A_0 \subset A$ tal que no existe un conjunto conexo en A que contenga a A_0 y que sea distinto de A_0. Así pues, vemos que *una componente conexa es un subconjunto conexo maximal*. Podemos definir una ***componente conexa por arcos*** de manera análoga, utilizando la conexión por arcos en vez de la conexión. Daremos algunas propiedades de las componentes conexas al final del capítulo.

3.5.3 Ejemplo *¿Es* $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, 3, \ldots\} \subset \mathbb{R}$ *conexo?*

Solución No, ya que si $U =]1/2, \infty[$ y $V =]-\infty, 1/4[$, entonces $\mathbb{Z} \subset U \cup V$, $\mathbb{Z} \cap U = \{1, 2, 3, \ldots\} \neq \varnothing$, $\mathbb{Z} \cap V = \{\ldots, -2, -1, 0\} \neq \varnothing$ y $\mathbb{Z} \cap U \cap V = \varnothing$. Por lo tanto, $\mathbb{Z}$ es disconexo (es decir, no es conexo). También es evidente que $\mathbb{Z}$ no es conexo por arcos, pero este solo hecho no puede usarse para concluir que $\mathbb{Z}$ no es conexo. ♦

3.5.4 Ejemplo *¿Es* $\{(x, y) \in \mathbb{R}^2 \mid 0 < x^2 + y^2 \leq 1\}$ *conexo?*

Solución Como en el ejemplo **3.4.2d**, sabemos que este conjunto es conexo por arcos, por lo que el teorema **3.5.2** implica que es conexo. Demostrar esto directamente es más difícil. ♦

Ejercicios de §3.5

1. ¿Es $[0, 1] \cup]2, 3]$ conexo? Demuéstrese la veracidad o falsedad de esta afirmación.

2. ¿Es $\{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1\} \cup \{(x, 0) \mid 1 < x < 2\}$ conexo? Demuéstrese la veracidad o falsedad de esta afirmación.

3. Sea $A \subset \mathbb{R}^2$ conexo por arcos. Considérese A como un subconjunto del plano xy en $\mathbb{R}^3$ y muéstrese que A sigue siendo conexo por arcos. ¿Puede obtenerse un argumento similar si A es conexo?

4. Analícense las componentes conexas de

a. $[0, 1] \cup [2, 3] \subset \mathbb{R}$

b. $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\} \subset \mathbb{R}$

c. $\{x \in [0, 1] \mid x \text{ es racional}\} \subset \mathbb{R}$

Demostraciones de los teoremas del capítulo 3

3.1.3 Teorema de Bolzano-Weierstrass *Un subconjunto de un espacio métrico es compacto sii es secuencialmente compacto.*

Demostración Comenzamos con dos lemas.

Lema 1 *Un conjunto compacto $A \subset M$ es cerrado.*

Demostración Mostraremos que $M \backslash A$ es abierto. Sea $x \in M \backslash A$ y considérese la siguiente colección de conjuntos abiertos: $U_n = \{y \mid d(x, y) > 1/n\}$. Puesto que cada $y \in M$ tal que $y \neq x$ cumple $d(y, x) > 0$, y estará en algún U_n. Así, los conjuntos U_n recubren A, por lo que debe existir un subrecubrimiento finito. Uno de ellos tiene el índice más grande, digamos, U_N. Si $\varepsilon = 1/N$, entonces, por construcción, $D(x, 1/N) \subset M \backslash A$, de modo que $M \backslash A$ es abierto. ▼

Lema 2 *Si M es un espacio métrico compacto y $B \subset M$ es cerrado, entonces B es compacto.*

Demostración Sean $\{U_i\}$ un recubrimiento abierto de B y $V = M \backslash B$, de modo que V es abierto. Entonces $\{U_i, V\}$ es un recubrimiento abierto de M. Por lo tanto, M tiene un recubrimiento finito, digamos, $\{U_1, \ldots, U_N, V\}$. Entonces $\{U_1, \ldots, U_N\}$ es un recubrimiento abierto finito de B. ▼

Demostración de 3.1.3 Sea A compacto y supongamos que existe una sucesión $x_k \in A$ que *no* tiene subsucesiones convergentes. En particular, esto significa que x_k tiene infinitos puntos distintos, digamos, $y_1, y_2, \ldots$. Puesto que no hay subsucesiones convergentes, existirá una vecindad U_k de y_k que no contiene otros y_j. Esto es así porque si toda vecindad de y_k contuviera otro y_j, podríamos elegir las vecindades $D(y_k, 1/m)$, $m = 1, 2, \ldots$ y seleccionar una sucesión convergente a y_k. Afirmamos que el conjunto $\{y_1, y_2, \ldots\}$ es cerrado. De hecho, no tiene puntos de acumulación, pues por hipótesis no tiene subsucesiones convergentes. Aplicamos el lema **2** a $\{y_1, y_2, \ldots\}$ como subconjunto de A y tenemos que $\{y_1, y_2, \ldots\}$ es compacto. Pero $\{U_k\}$ es un recubrimiento abierto que no contiene un subrecubrimiento finito, lo cual es una contradicción. Así, x_k tiene una subsucesión convergente. El límite está en A, pues A es cerrado, por el lema **1**.

Recíprocamente, supóngase que A es secuencialmente compacto. Para demostrar que A es compacto, sea $\{U_i\}$ un recubrimiento abierto de A. Necesitamos demostrar que tiene un subrecubrimiento finito. Para ello procedemos siguiendo varios pasos.

Lema 3 *Existe $r > 0$ tal que para todo $y \in A$, $D(y, r) \subset U_i$ para algún U_i.*

Demostración En caso contrario, para todo n existirá y_n tal que $D(y_n, 1/n)$ no está contenido en ningún U_i. Por hipótesis, y_n tiene una subsucesión convergente, digamos, $z_n \to z \in A$. Como los conjuntos U_i recubren A, $z \in U_{i_0}$ para algún U_{i_0}. Elijamos $\varepsilon > 0$ tal que $D(z, \varepsilon) \subset U_{i_0}$, lo que es posible porque U_{i_0} es abierto. Sea N lo suficientemente grande para que $d(z_N, z) < \varepsilon/2$ y $1/N < \varepsilon/2$. Entonces $D(z_N, 1/N) \subset U_{i_0}$, una contradicción. ▼

Lema 4 *A es totalmente acotado (véase la definición **3.1.4**).*

Demostración Si A no fuese totalmente acotado, entonces para algún $\varepsilon > 0$ no podríamos recubrir A con un número finito de discos. Elijamos $y_1 \in A$ e $y_2 \in A \backslash D(y_1, \varepsilon)$. Por hipótesis se puede repetir el proceso: elijamos $y_n \in A \backslash [D(y_1, \varepsilon) \cup \cdots \cup D(y_{n-1}, \varepsilon)]$. Ésta sería una sucesión tal que $d(y_n, y_m) \geq \varepsilon$ para todo n y m, por lo que y_n no tendría ninguna subsucesión convergente, lo que contradice la hipótesis de que A es secuencialmente compacto. ▼

Para terminar nuestra demostración, sea r como en el lema **3**. Por el lema **4**, podemos escribir $A \subset D(y_1, r) \cup \cdots \cup D(y_n, r)$ para un número finito de y_j. Por el lema **3**, $D(y_j, r) \subset U_{i_j}$, $j = 1, \ldots, n$ para algún índice i_j. Entonces $U_{i_1}, \ldots, U_{i_n}$ recubren A. ■

3.1.5 Teorema *Un espacio métrico es compacto sii es completo y totalmente acotado.*

Demostración Primero supongamos que M es compacto. Por **3.1.3**, es secuencialmente compacto. Así, si x_k es una sucesión de Cauchy, tiene una subsucesión convergente y entonces, como en **1.4.7**, toda la sucesión converge. Por lo tanto, M es completo. También es totalmente acotado, por el lema **4**.

Recíprocamente, supóngase que M es completo y totalmente acotado. Por **3.1.3**, basta mostrar que M es secuencialmente compacto. Sea y_k una sucesión en M; podemos suponer que los términos y_k son todos distintos, ya que si y_k se repite infinitas veces, existe una subsucesión convergente trivial, y si existe un número finito de repeticiones, podemos simplemente eliminarlas. Dado un entero N, recubrimos M con un número finito de bolas $D(x_{L_1}, 1/N), \ldots, D(x_{L_N}, 1/N)$. En una de estas bolas hay infinitos y_k. Comencemos con $N = 1$. Escribamos $M = D(x_{L_1}, 1) \cup \cdots \cup D(x_{L_N}, 1)$, de modo que podemos seleccionar una subsucesión de y_k que está totalmente contenida en una de estas bolas. Repitamos el procedimiento para $N = 2$, para obtener otra subsucesión contenida en una bola fija de radio $1/2$, y así sucesivamente. Ahora elijamos la subsucesión "diagonal":

el primer elemento de la primera sucesión, el segundo de la segunda, etcétera. Esta sucesión es de Cauchy y, como M es completo, converge. ■

3.2.1 Teorema de Heine-Borel *Un conjunto $A \subset \mathbb{R}^n$ es compacto sii es cerrado y acotado.*

Demostración Ya hemos demostrado que los conjuntos compactos son cerrados y acotados; ahora debemos mostrar que un conjunto $A \subset \mathbb{R}^n$ es compacto si es cerrado y acotado. Daremos dos demostraciones de este hecho.

Primera demostración Esta demostración se basa en el teorema de Bolzano-Weierstrass y el hecho de que cualquier sucesión acotada en $\mathbb{R}$ tiene una subsucesión convergente, que demostramos en **1.4.3.** De hecho, demostraremos que un conjunto cerrado y acotado A es secuencialmente compacto. Sea $x_k = (x_k^1, x_k^2, \ldots, x_k^n) \in \mathbb{R}^n$ una sucesión. Como A está acotado, x_k^1 tiene una subsucesión convergente, digamos, $x_{f_1(k)}^1$. Entonces $x_{f_1(k)}^2$ tiene una subsucesión convergente, digamos $x_{f_2(k)}^2$. Continuamos de esta manera para obtener una subsucesión $x_{f_n(k)} = ((x_{f_n(k)}^1, \ldots, x^n_{f_n(k)})$, tal que todas sus componentes convergen. Así, $x_{f_n(k)}$ converge en $\mathbb{R}^n$. El límite está en A, puesto que A es cerrado. Así, A es secuencialmente compacto, y por tanto compacto. ■

Segunda demostración Esta demostración usa directamente la definición de compacidad en términos de recubrimientos abiertos. Comenzamos con un caso particular:

Lema 1 *Los intervalos cerrados $[a, b]$ en $\mathbb{R}$ son compactos.*

Demostración Sea $U = \{U_i\}$ un recubrimiento abierto de $[a, b]$. Defínase

$$C = \{x \in [a, b] \mid \text{el conjunto } [a, x] \text{ puede recubrirse con una colección finita de los } U_i\}.$$

Queremos demostrar que $C = [a, b]$. Para esto, sea $c = \sup(C)$. El supremo existe, pues $C \neq \varnothing$ (ya que $a \in C$) y C está acotado superiormente por b. Como $a \in C$ y b es una cota superior de C, $c \in [a, b]$, por definición de $\sup(C)$. Supongamos que $c \in U_{i_0}$; tal U_{i_0} existe porque los conjuntos U_i recubren $[a, b]$. Como U_{i_0} es abierto, existe $\varepsilon > 0$ tal que $]c - \varepsilon, c + \varepsilon[\subset U_{i_0}$. Como $c = \sup(C)$, existe $x \in C$ tal que $c - \varepsilon < x \leq c$ (véase la proposición **1.3.2**). Como $x \in C$, $[a, x]$ tiene un subrecubrimiento finito, digamos, $U_1, \ldots, U_N$; entonces $[a, c + \varepsilon/2]$ también tiene el subrecubrimiento finito $U_1, \ldots, U_N$, U_{i_0}. Así, podemos concluir que $c \in C$ y, además, que $c = b$. De hecho, si $c < b$, obtendríamos un elemento de C mayor que c, pues $[a, c + \varepsilon/2]$ tiene un subrecubrimiento finito, lo cual no puede ocurrir, ya que $c = \sup(C)$. ▼

Nota. ¿Por qué no sirve esta demostración para $]a, b]$, $[a, b[$ o $[a, \infty[$?

Lema 2 *Si $A \subset \mathbb{R}^n$ es compacto y $x_0 \in \mathbb{R}^m$, entonces $A \times \{x_0\} \subset \mathbb{R}^n \times \mathbb{R}^m$ es compacto.*

Demostración Sea $\mathcal{U}$ un recubrimiento abierto de $A \times \{x_0\}$ y

$$\mathcal{V} = \{V \mid V = \{y \mid (y, x_0) \in U\}, \text{ para algún } U \in \mathcal{U}\}.$$

Entonces $\mathcal{V}$ es un recubrimiento abierto de A en $\mathbb{R}^n$, y por lo tanto, $\mathcal{V}$ contiene un subrecubrimiento finito de A, digamos, $\mathcal{V}' = \{V_1, \ldots, V_k\}$. Cada $V_i \in \mathcal{V}'$ corresponde a un $U_i \in \mathcal{U}$; $\mathcal{U}' = \{U_1, \ldots, U_k\}$ es entonces un subrecubrimiento finito en $\mathbb{R}^n \times \mathbb{R}^m$ de $A \times \{x_0\}$. ▼

El siguiente paso es un argumento de inducción.

Lema 3 *Si $[-R, R]^{n-1} \subset \mathbb{R}^{n-1}$ es compacto, entonces $[-R, R]^n \subset \mathbb{R}^n$ es compacto, donde $[-R, R]^n = [-R, R] \times \cdots \times [-R, R]$, n veces.*

Demostración Supóngase que $[-R, R]^{n-1}$ es compacto y que $\mathcal{U}$ es un recubrimiento abierto de $[-R, R]^n$. Definamos

$$S = \{x \in [-R, R] \mid [-R, R]^{n-1} \times [-R, x] \subset \mathbb{R}^n \text{ tiene un subrecubrimiento finito en } \mathcal{U}\}.$$

Ahora, $-R \in S$, pues $[-R, R]^{n-1}$ es compacto, por hipótesis; entonces, el lema **2** implica que $[-R, R]^{n-1} \times \{-R\}$ tiene un subrecubrimiento finito en U. Como S está acotado superiormente por R, tiene un supremo, digamos, x_0. Probaremos que $x_0 = R$, lo que demostrará el lema.

Sea $\mathcal{U}' \subset \mathcal{U}$ un subrecubrimiento finito de $[-R, R]^{n-1} \times \{x_0\}$. Para cada punto $(y, x_0) \in [-R, R]^{n-1} \times \{x_0\}$, existe $\varepsilon_y > 0$ tal que $D((y, x_0), \sqrt{2}\varepsilon_y)$ está recubierto por $\mathcal{U}'$. Como

$$V_y = D(y, \varepsilon_y) \times]x_0 - \varepsilon_y, x_0 + \varepsilon_y[\subset D((y, x_0), \sqrt{2}\varepsilon_y),$$

éste estará recubierto por $\mathcal{U}'$. Consideremos el recubrimiento abierto $\mathcal{V} = \{V_y \mid y \in [-R, R]^{n-1}\}$ de $[-R, R]^{n-1} \times \{x_0\}$. Por el lema **2**, $\mathcal{V}$ contiene un subrecubrimiento finito de $[-R, R]^{n-1} \times \{x_0\}$, digamos $\{V_{y_1}, \ldots, V_{y_N}\}$. Sea $\varepsilon = \inf\{\varepsilon_{y_1}, \ldots, \varepsilon_{y_N}\}$. Entonces

$$[-R, R]^{n-1} \times]x_0 - \varepsilon, x_0 + \varepsilon[\subset \bigcup_{i=1}^{\infty} V_{y_i},$$

y, por lo tanto, $[-R, R]^{n-1} \times]x_0 - \varepsilon, x_0 + \varepsilon[$ está recubierto por $\mathcal{U}'$.

Con este ε, existirá $x \in S$ tal que $x_0 - \varepsilon < x \leq x_0$. Como $x \in S$, existe un subrecubrimiento finito $U'' \subset U$ que recubre $[-R, R]^{n-1} \times [-R, x]$, y $U' \cup U''$ es un recubrimiento finito de $[-R, R]^{n-1} \times [-R, x_0 + \varepsilon[$. Así, $x_0 \in S$. Supongamos que $x_0 < R$; sea δ tal que $x_0 + \delta < R$ y $x_0 + \delta < x_0 + \varepsilon$. Así, $[-R, R]^{n-1} \times [-R, x_0 + \delta]$ está recubierto por $U' \cup U''$, y $x_0 + \delta \in S$, una contradicción; en consecuencia, $x_0 = R$. ▼

Para concluir la demostración del teorema de Heine-Borel, sea $A \subset \mathbb{R}^n$ cerrado y acotado. Como está acotado, existe $R > 0$ tal que $A \subset [-R, R]^n$. Los lemas **1** y **3** demuestran que $[-R, R]^n$ es compacto. El lema **2** en la demostración de **3.1.3** demuestra que A es compacto, pues es un subespacio cerrado del conjunto compacto $[-R, R]^n$. ■

3.3.1 Propiedad de los conjuntos encajados *Sea F_k una sucesión de conjuntos compactos no vacíos en un espacio métrico M tales que $F_{k+1} \subset F_k$ para todo $k =$ 1, 2, ... Entonces existe al menos un punto en $\bigcap_{i=1}^{\infty} F_k$.*

Demostración En el conjunto compacto $A = F_1$, los conjuntos $F_1, F_2, \ldots$ tienen la propiedad de intersección finita, pues la intersección de cualquier colección finita de ellos es igual al conjunto F_k de índice mayor. Por el ejemplo **3.1.9,**

$$F_1 \bigcap \left(\bigcap_{k=1}^{\infty} F_k \right) = \bigcap_{k=1}^{\infty} \{F_k\} \neq \varnothing. \quad ■$$

3.5.2 Teorema *Los conjuntos conexos por arcos son conexos.*

Comenzaremos demostrando un caso particular del teorema.

Lema *El intervalo $[a, b]$ es conexo.*

Demostración Supongamos que el intervalo no fuera conexo. Entonces existirían dos conjuntos abiertos U y V tales que $U \cap [a, b] \neq \varnothing$ y $V \cap [a, b] \neq \varnothing$, $[a, b] \cap U \cap V = \varnothing$, y $[a, b] \subset U \cup V$. Supongamos, además, que $b \in V$. Sea $c = \sup(U \cap [a, b])$, que existe porque $U \cap [a, b]$ es no vacío y está acotado superiormente. El conjunto $U \cap [a, b]$ es cerrado, pues su complementario es $V \cup (\mathbb{R} \backslash [a, b])$, que es abierto. Así, $c \in U \cap [a, b]$ (véase el ejercicio 8 del capítulo 2). Ahora, $c \neq b$, pues $c \notin V$ y $b \in V$. Afirmamos que cualquier vecindad de c interseca a $V \cap [a, b]$. Para verlo, obsérvese que

$c \neq b$ y que ninguna vecindad de c puede estar totalmente contenida en U, pues $c = \sup(U \cap [a, b])$. Esto demuestra la afirmación y muestra que c es un punto de acumulación de $V \cap [a, b]$. Como $U \cap [a, b]$, el conjunto $V \cap [a, b]$ es cerrado, y entonces $c \in V \cap [a, b]$. Esto contradice la afirmación de que $V \cap U \cap [a, b] = \varnothing$. ▼

Corolario *No existen subconjuntos cerrados disjuntos no vacíos C, D de $[a, b]$ cuya unión sea $[a, b]$.*

Esto se sigue del lema, si $U = \mathbb{R} \backslash D$ y $V = \mathbb{R} \backslash C$.

Demostración del teorema 3.5.2 Supongamos que A no es conexo. Por definición, existen dos conjuntos abiertos U, V que separan A; es decir, los conjuntos U y V son tales que $A \subset U \cup V$, $A \cap U \cap V = \varnothing$, $U \cap A \neq \varnothing$, y $V \cap A \neq \varnothing$.

Sean $x \in U \cap A$ e $y \in V \cap A$. Como A es conexo por arcos, existe un arco continuo $\varphi : [a, b] \to A$ que une x con y. Sean $C = \varphi^{-1}(U)$ y $D = \varphi^{-1}(V)$, de modo que $C, D \subset [a, b]$. Supongamos que $t_k \in C$ y que $t_k \to t$. Entonces $t \in [a, b]$, pues $[a, b]$ es cerrado; $\varphi(t_k) \in U$, por definición de C, y $\varphi(t_k) \to \varphi(t)$, por continuidad. Afirmamos que $\varphi(t) \in U$; en caso contrario, $\varphi(t) \in V$ y como V es abierto, $\varphi(t_k) \in V$ para k suficientemente grande, lo que contradice el hecho de que $\varphi(t_k) \in U$, que no interseca a $A \cap V$. Así, $\varphi(t) \in U$ y entonces C es cerrado. Análogamente, D es cerrado. Entonces C, D son no vacíos, pues $a \in C$ y $b \in D$, y son disjuntos, ya que $C \cap D = \varphi^{-1}(U) \cap \varphi^{-1}(V) = \varphi^{-1}(U \cap V) = \varnothing$ y $C \cup D = \varphi^{-1}(U) \cup \varphi^{-1}(V) = \varphi^{-1}(U \cup V) = \varphi^{-1}(A) = [a, b]$. Obtenemos, pues, una contradicción con el corolario del lema, luego A debe ser conexo. ■

Ejemplos resueltos del capítulo 3

Ejemplo 3.1 *Muéstrese que $A = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$ es compacto y conexo.*

Solución Para mostrar que A es compacto, mostraremos que es cerrado y acotado. Para mostrar que es cerrado, consideremos $\mathbb{R}^n \backslash A = \{x \in \mathbb{R}^n \mid \|x\| > 1\} = B$. Para todo $x \in B$, $D(x, \|x\| - 1) \subset B$, de modo que B es abierto y, por tanto, A es cerrado. Claramente A está acotado, pues $A \subset D(0, 2)$; por lo tanto, A es compacto.

Para mostrar que A es conexo, mostraremos que es conexo por arcos. Sean $x, y \in A$. Entonces, la línea recta que une x e y es el arco pedido. Explícitamente, usamos $\varphi : [0, 1] \to \mathbb{R}^n$, $\varphi(t) = (1 - t)x + ty$. Se puede ver que $\varphi(t) \in A$, pues $\|\varphi(t)\| \leq (1 - t)\|x\| + t\|y\| \leq (1 - t) + t = 1$, por la desigualdad triangular. ♦

Ejemplo 3.2 *Sean $A \subset \mathbb{R}^n$, $x \in A$ e $y \in \mathbb{R}^n \backslash A$. Sea $\varphi: [0, 1] \to \mathbb{R}^n$ un arco (continuo) que une x e y. Muéstrese que existe t tal que $\varphi(t) \in \partial A$.*

Solución Intuitivamente, este resultado dice que un arco que une un conjunto con su complemento debe atravesar la frontera del conjunto en algún punto. Véase la figura 3.E-1.

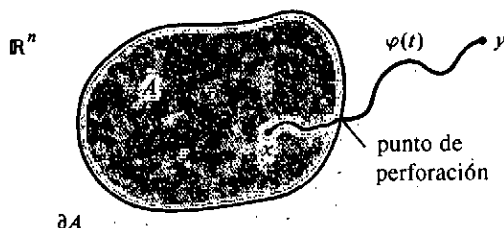


FIGURA 3.E-1 Un arco que une A con $\mathbb{R}^n \backslash A$

Sea $B = \{x \in [0, 1] \mid \varphi([0, x]) \subset A\} \subset [0, 1]$. Entonces, $B \neq \varnothing$, pues $0 \in B$. Sea $t = \sup(B)$. Mostraremos que $\varphi(t) \in \partial A$. Sea U una vecindad de $\varphi(t)$. Elijamos $t_k \in [0, t]$, $t_k \to t$, tal que $\varphi(t_k) \in A$. Esto es posible, por la definición de B. Entonces $\varphi(t_k) \in U$ para k suficientemente grande, por la continuidad de φ. Por la definición de t, existe un punto s_k tal que $t \leq s_k \leq t + 1/k$ y $\varphi(s_k) \in A$. Ahora, $s_k \to t$ y entonces, por la continuidad de φ, $\varphi(s_k) \notin U$ para k suficientemente grande. Así, U contiene puntos de A y de $\mathbb{R}^n \backslash A$; por la proposición **2.6.2**, $\varphi(t) \in \partial A$. ♦

Ejemplo 3.3 *Demuéstrese que un conjunto $A \subset \mathbb{R}$ es conexo sii es un intervalo (un intervalo es un conjunto de la forma $[a, b]$, $[a, b[$, $]a, b]$ o $]a, b[$, donde a o b pueden ser $\pm\infty$ en un extremo abierto del intervalo).*

Solución Ya hemos visto que los intervalos son conexos, pues son conexos por arcos. Ahora, para el recíproco, supóngase que A no es un intervalo. Mostraremos que entonces no es conexo. Decir que A no es un intervalo significa que existen puntos x, y, z tales que $x < y < z$, con $x, z \in A$ e $y \notin A$ (¿por qué?). Entonces $U =]-\infty, y[$ y $V =]y, \infty[$ son conjuntos abiertos tales que $A \subset U \cup V$, $U \cap V \cap A = \varnothing$, $U \cap A \neq \varnothing$ y $V \cap A \neq \varnothing$. Por lo tanto, A no es conexo. ♦

Ejemplo 3.4 *Sea A un subconjunto conexo y abierto de $\mathbb{R}^n$. Demuéstrese que A es conexo por arcos.*

Solución La solución ilustra una técnica importante para el uso de la conexión. Sean $x_0 \in A$ y $B = \{y \in A \mid x_0$ e y se puede unir mediante un arco continuo$\}$. Obviamente $x_0 \in B$, de modo que $B \neq \varnothing$. Afirmamos que B es abierto y cerrado como subconjunto de A; es decir, B es abierto y cerrado con respecto de A. En primer lugar, B es abierto, por la siguiente razón: si $y \in B$, elijamos un disco $D(y, \varepsilon) \subset A$, lo que es posible ya que A es abierto. Si $z \in D(y, \varepsilon)$, entonces $z \in B$, pues podemos construir un arco continuo de x_0 a z mediante la concatenación de un arco de x_0 a y con la línea recta de y a z (el lector debe demostrar que esto produce un arco *continuo*). Así, $D(y, \varepsilon) \subset B$, de modo que B es abierto. Para probar que B es cerrado, sea $y_k \in B$ tal que $y_k \to y \in A$. Como A es abierto, existe $\varepsilon > 0$ tal que $D(y, \varepsilon) \subset A$. Como $y_k \to y$, existe N tal que $y_k \in D(y, \varepsilon)$ para $k \geq N$. Al unir x_0 con y_N mediante un arco continuo, seguido de la línea recta de y_N a y, vemos que $y \in B$, por lo que B es cerrado. Como $B \neq \varnothing$ y B es abierto y cerrado en A, obtenemos $B = A$ (en caso contrario, B y $B \backslash A$ separarían a A). Así, todo punto en A se puede unir con x_0 mediante un arco continuo, luego A es conexo por arcos. ♦

Ejercicios del capítulo 3

1. ¿Cuáles de los siguientes conjuntos son compactos? ¿Cuáles son conexos?

- **a.** $\{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| \leq 1\}$
- **b.** $\{x \in \mathbb{R}^n \mid \|x\| \leq 10\}$
- **c.** $\{x \in \mathbb{R}^n \mid 1 \leq \|x\| \leq 2\}$
- **d.** $\mathbb{Z} = \{\text{enteros en } \mathbb{R}\}$
- **e.** Un conjunto finito en $\mathbb{R}$
- **f.** $\{x \in \mathbb{R}^n \mid \|x\| = 1\}$ (distíngase entre los casos $n = 1$ y $n \geq 2$)
- **g.** El perímetro del cuadrado unidad en $\mathbb{R}^2$
- **h.** La frontera de un conjunto acotado en $\mathbb{R}$
- **i.** Los racionales en $[0, 1]$
- **j.** Un conjunto cerrado en $[0, 1]$

2. Demuéstrese que un conjunto $A \subset \mathbb{R}^n$ no es conexo sii podemos escribir $A \subset F_1 \cup F_2$, donde F_1, F_2 son cerrados, $A \cap F_1 \cap F_2 = \varnothing$, $F_1 \cap A \neq \varnothing$, $F_2 \cap A \neq \varnothing$.

3. Demuéstrese que en $\mathbb{R}^n$ un conjunto infinito acotado A tiene un punto de acumulación.

4. Muéstrese que un conjunto A está acotado sii existe una constante M tal que $d(x, y) \leq M$ para todo $x, y \in A$. Proporciónese una definición plausible del diámetro de un conjunto y reformúlese su resultado.

5. Muéstrese que los siguientes conjuntos no son compactos, exhibiendo un recubrimiento abierto que no contenga un subrecubrimiento finito.

a. $\{x \in \mathbb{R}^n \mid \|x\| < 1\}$.

b. $\mathbb{Z}$, los enteros en $\mathbb{R}$

6. Supóngase que F_k es una sucesión de conjuntos compactos no vacíos que satisfacen la propiedad de los conjuntos encajados, de modo que diámetro$(F_k) \to 0$ cuando $k \to \infty$. Muéstrese que existe exactamente un punto en $\cap \{F_k\}$. (Por definición, diámetro $(F_k) = \sup\{d(x, y) \mid x, y \in F_k\}$).

7. Sea x_k una sucesión en $\mathbb{R}^n$ que converge a x y sean $A_k = \{x_k, x_{k+1}, \ldots\}$. Muéstrese que $\{x\} = \cap_{k=1}^{\infty} \mathrm{cl}(A_k)$. ¿Es esto cierto para *cualquier* espacio métrico?

8. Sean $A \subset \mathbb{R}^n$ un conjunto compacto y x_k una sucesión de Cauchy en $\mathbb{R}^n$ tal que $x_k \in A$. Muéstrese que x_k converge a un punto en A.

9. Determínese (mediante una demostración o un contraejemplo) la veracidad o falsedad de los siguientes enunciados:

a. (A es compacto en $\mathbb{R}^n$) $\Rightarrow$ ($\mathbb{R}^n \backslash A$ es conexo).

b. (A es conexo en $\mathbb{R}^n$) $\Rightarrow$ ($\mathbb{R}^n \backslash A$ es conexo).

c. (A es conexo en $\mathbb{R}^n$) $\Rightarrow$ (A es abierto o cerrado).

d. ($A = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$) $\Rightarrow$ ($\mathbb{R}^n \backslash A$ es conexo) [Sugerencia: verifíquense los casos $n = 1$ y $n \geq 2$.]

10. Un espacio métrico M es ***localmente conexo por arcos*** si cada punto de M tiene una vecindad U tal que U es conexo por arcos. (Esta terminología difiere un poco de la que aparece en los libros de topología.) Muéstrese que (M es conexo y localmente conexo por arcos) $\Leftrightarrow$ (M es conexo por arcos).

11. **a.** Demuéstrese que si A es conexo en un espacio métrico M y $A \subset B \subset \mathrm{cl}(A)$, entonces B es conexo.

b. Dedúzcase de **a** que las componentes conexas de un conjunto A son relativamente cerradas. Encuéntrese un ejemplo en el que no sean relativamente abiertas. ($C \subset A$ es ***relativamente cerrado*** en A si C es la intersección de algún conjunto cerrado en M con A; es decir, si C es cerrado en el espacio métrico A.)

c. Muéstrese que si una familia $\{B_i\}$ de conjuntos conexos es tal que $B_i \cap B_j \neq \varnothing$ para todo i, j, entonces $\cup_i B_i$ es conexo.

d. Dedúzcase de **c** que todo punto de un conjunto está dentro de una única componente conexa.

e. Utilícese **c** para mostrar que $\mathbb{R}^n$ es conexo, partiendo del hecho de que las rectas en $\mathbb{R}^n$ son conexas.

12. Sea S un conjunto de números reales que no es vacío y que está acotado superiormente. Sea $-S = \{x \in \mathbb{R} \mid -x \in S\}$. Demuéstrese que

a. $-S$ está acotado inferiormente.

b. $\sup S = -\inf(-S)$.

13. Sea M un espacio métrico completo y F_n una colección de subconjuntos cerrados no vacíos (no necesariamente compactos) de M tales que $F_{n+1} \subset F_n$ y diámetro $(F_n) \to 0$. Demuéstrese que $\bigcap_{n=1}^{\infty} F_n$ consta de un único punto; compárese este resultado con el del ejercicio 6.

14. **a.** Un punto $x \in A \subset M$ es ***aislado*** en el conjunto A si existe una vecindad U de x tal que $U \cap A = \{x\}$. Muéstrese que esto es equivalente a decir que existe $\varepsilon > 0$ tal que para todo $y \in A$, $y \neq x$, tenemos $d(x, y) > \varepsilon$.

b. Un conjunto es ***discreto*** si todos sus puntos son aislados. Proporciónense algunos ejemplos. Muéstrese que un conjunto discreto es compacto sii es finito.

15. Sean $K_1 \subset M_1$ y $K_2 \subset M_2$ conjuntos conexos por arcos (respectivamente, conexos, compactos). Muéstrese que $K_1 \times K_2$ es conexo por arcos (respectivamente, conexo, compacto) en $M_1 \times M_2$.

16. Si $x_k \to x$ en un espacio normado, demuéstrese que $\|x_k\| \to \|x\|$. ¿Es cierto el recíproco? Utilícese este resultado para demostrar que $\{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$ es cerrado por medio de sucesiones.

17. Sea K un conjunto cerrado no vacío en $\mathbb{R}^n$ y $x \in \mathbb{R}^n \backslash K$. Demuéstrese que existe $y \in K$ tal que $d(x, y) = \inf\{d(x, z) \mid z \in K\}$ ¿Es cierto esto para conjuntos abiertos? ¿Es cierto en espacios métricos generales?

18. Sea $F_n \subset \mathbb{R}$ dado por $F_n = \{x \mid x \geq 0 \text{ y } 2 - 1/n \leq x^2 \leq 2 + 1/n\}$. Muéstrese que $\bigcap_{n=1}^{\infty} F_n \neq \varnothing$. Utilícese este resultado para probar la existencia de $\sqrt{2}$.

19. Sean $V_n \subset M$ conjuntos abiertos tales que $\mathrm{cl}(V_n)$ es compacto, $V_n \neq \varnothing$ y $\mathrm{cl}(V_n) \subset V_{n-1}$. Demuéstrese que $\bigcap_{n=1}^{\infty} V_n \neq \varnothing$.

20. Demuéstrese que un conjunto compacto A es cerrado de la siguiente forma: sea x un punto de acumulación de A y supongamos que $x \notin A$; para cada $y \in A$ elegimos vecindades disjuntas U_y de y y V_y de x. Considérese el recubrimiento abierto $\{U_y\}$.

21. **a.** Demuéstrese que un conjunto $A \subset M$ es conexo sii $\varnothing$ y A son los únicos subconjuntos de A que son relativamente abiertos y cerrados en A. (Un conjunto $U \subset A$ es ***relativamente abierto*** en A si $U = V \cap A$ para algún conjunto abierto $V \subset M$; "relativamente cerrado en A" se define de manera análoga.)

b. Demuéstrese que $\varnothing$ y $\mathbb{R}^n$ son los únicos subconjuntos de $\mathbb{R}^n$ que son abiertos y cerrados a la vez.

22. Encuéntrense dos subconjuntos $A, B \subset \mathbb{R}^2$ y un punto $x_0 \in \mathbb{R}^2$ tales que $A \cup B$ no sea conexo pero que $A \cup B \cup \{x_0\}$ sí lo sea.

23. Sean $\mathbb{Q}$ los racionales en $\mathbb{R}$. Muéstrese que ni $\mathbb{Q}$ ni los irracionales $\mathbb{R}\backslash\mathbb{Q}$ son conexos.

24. Demuéstrese que un conjunto $A \subset M$ no es conexo si podemos escribir A como la unión disjunta de dos conjuntos B y C de modo que $B \cap A \neq \varnothing$, $C \cap A \neq \varnothing$, y que ninguno de los conjuntos B o C tenga un punto de acumulación que pertenezca al otro conjunto.

25. Demuéstrese que existe una sucesión de enteros distintos $n_1, n_2, \ldots \to \infty$ tales que $\lim_{k\to\infty} \operatorname{sen} n_k$ existe.

26. Muéstrese que puede reemplazarse la propiedad de completitud de $\mathbb{R}$ por la propiedad de los conjuntos encajados.

27. Sea $A \subset \mathbb{R}$ un conjunto acotado. Muéstrese que A es cerrado sii para toda sucesión $x_n \in A$, $\lim\sup x_n \in A$ y $\lim\inf x_n \in A$.

28. Sea $A \subset M$ un conjunto conexo con más de un punto. Muéstrese que cada punto de A es un punto de acumulación de A.

29. Sea $A = \{(x, y) \in \mathbb{R}^2 \mid x^4 + y^4 = 1\}$. Muéstrese que A es compacto. ¿Es conexo?

30. Sea U_k una sucesión de conjuntos abiertos acotados en $\mathbb{R}^n$. Demuéstrese o refútese:

a. $\bigcup_{k=1}^{\infty} U_k$ es abierto.

b. $\bigcap_{k=1}^{\infty} U_k$ es abierto.

c. $\bigcap_{k=1}^{\infty} (\mathbb{R}^n\backslash U_k)$ es cerrado.

d. $\bigcap_{k=1}^{\infty} (\mathbb{R}^n\backslash U_k)$ es compacto.

31. Supongamos que $A \subset \mathbb{R}^n$ no es compacto. Muéstrese que existe una sucesión $F_1 \supset F_2 \supset F_3 \cdots$ de conjuntos cerrados tales que $F_k \cap A \neq \varnothing$ para todo k y que

$$\left(\bigcap_{k=1}^{\infty} F_k\right) \bigcap A = \varnothing.$$

32. Sea x_n una sucesión en $\mathbb{R}^3$ tal que $\|x_{n+1} - x_n\| \leq 1/(n^2 + n)$, $n \geq 1$. Muéstrese que x_n converge.

33. ***Teorema de la categoría de Baire.*** Un conjunto S en un espacio métrico es ***nunca denso*** si para cada conjunto abierto no vacío U, tenemos $\operatorname{cl}(S) \cap U \neq U$, o, de forma equivalente, $\operatorname{int}(\operatorname{cl}(S)) \neq \varnothing$. Muéstrese que $\mathbb{R}^n$ no se puede escribir como la unión numerable de conjuntos nunca densos.

34. Demuéstrese que cualquier conjunto cerrado $A \subset M$ es la intersección de una familia numerable de conjuntos abiertos.

35. Sea $a \in \mathbb{R}$ y definamos la sucesión $a_1, a_2, \ldots$ en $\mathbb{R}$ como $a_1 = a$ y $a_n = a^2_{n-1} - a_{n-1} + 1$ si $n > 1$. ¿Para qué valores de $a \in \mathbb{R}$ la sucesión es

a. monótona?

b. acotada?

c. convergente?

Calcúlese el límite en los casos en que converja.

36. Sea $A \subset \mathbb{R}^n$ no numerable. Demuéstrese que A tiene un punto de acumulación.

37. Sean $A, B \subset M$ con A compacto, B cerrado y $A \cap B = \varnothing$.

a. Muéstrese que existe $\varepsilon > 0$ tal que $d(x, y) > \varepsilon$ para todo $x \in A$ e $y \in B$.

b. ¿Es cierto esto si A, B son cerrados solamente?

38. Muéstrese que $A \subset M$ no es conexo sii existen dos conjuntos abiertos *disjuntos* U, V tales que $U \cap A \neq \varnothing$, $V \cap A \neq \varnothing$ y $A \subset U \cup V$.

39. Sea $F_1 = [0, 1/3] \cup [2/3, 1]$, obtenido de $[0, 1]$ eliminando el tercio del medio. Repitiendo el proceso obtenemos

$$F_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1].$$

En general, F_n es una unión de intervalos y F_{n+1} se obtiene al eliminar el tercio del medio de cada uno de estos intervalos. Sea $C = \bigcap_{n=1}^{\infty} F_n$, el ***conjunto de Cantor***. Demuéstrese:

a. C es compacto.

b. C tiene infinitos puntos. [Sugerencia: analícense los extremos de F_n.]

c. $\text{int}(C) = \varnothing$.

d. C es ***perfecto***; es decir, es cerrado y no tiene puntos aislados.

e. Muéstrese que C es ***totalmente disconexo***; es decir, si $x, y \in C$ y $x \neq y$, entonces $x \in U$ e $y \in V$, donde U y V son conjuntos abiertos que separan C.

40. Sea F_k una colección de conjuntos compactos encajados (es decir, $F_{k+1} \subset F_k$). Supongamos, además, que cada F_k es conexo. Demuéstrese que $\bigcap_{k=1}^{\infty} \{F_k\}$ es conexo. Encuéntrese un ejemplo que muestre que la compacidad es una condición esencial y que no podemos suponer simplemente que "F_k es un encaje de conjuntos conexos cerrados".

Capítulo 4
Transformaciones continuas

Para obtener teoremas útiles e interesantes, a veces es necesario imponer restricciones sobre los objetos matemáticos que se estudian. En este capítulo imponemos la condición de que las funciones analizadas sean continuas, e investigaremos algunas consecuencias de esta restricción. En el capítulo 6 estudiaremos una restricción aún mayor: la diferenciabilidad.

§4.1 Continuidad

En primer lugar analizaremos el concepto de continuidad para las funciones de valores reales sobre la recta real $\mathbb{R}$. La figura 4.1-1a muestra una función continua y la figura 4.1-1b muestra una función discontinua. Una función continua tiene la importante propiedad de que si x está cerca de x_0, $f(x)$ está cerca de $f(x_0)$ (como lo muestra la figura 4.1-1a). Por otro lado, en la figura 4.1-1b, aunque x esté muy cerca de x_0, $f(x)$ no necesariamente está cerca de $f(x_0)$. El lector debería estar familiarizado con estas ideas de sus cursos de cálculo de una variable.

Para definir la continuidad en términos más generales y precisos, primero vamos a definir el concepto de límite de una función en un punto. Sean (M, d) y (N, ρ) dos espacios métricos, $A \subset M$, y $f\colon A \to N$ una función dada.

4.1.1 Definición *Supongamos que x_0 es un punto de acumulación de A. Decimos que $b \in N$ es el* ***límite de f en*** *x_0, lo que denotamos*

$$\lim_{x \to x_0} f(x) = b,$$

si dado $\varepsilon > 0$ existe $\delta > 0$ (que posiblemente depende de f, x_0 y ε) tal que para todo $x \in A$ que satisfaga $x \neq x_0$ y $d(x_0, x)$, tenemos que $\rho(f(x), b) < \varepsilon$.

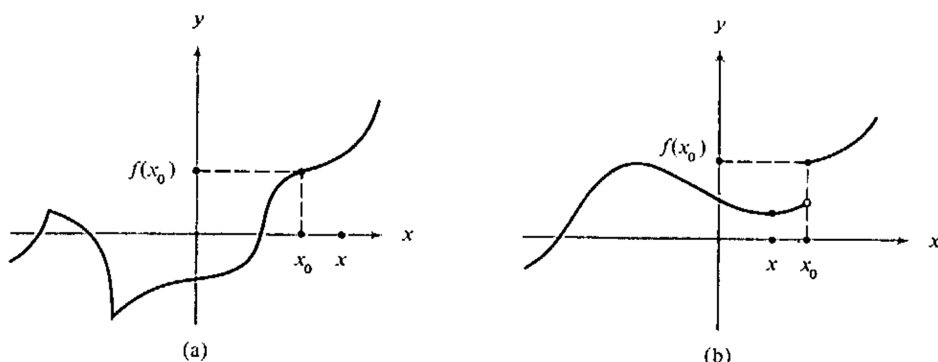


FIGURA 4.1-1 (a) Función continua; (b) función discontinua

Intuitivamente, esto quiere decir que cuando x tiende a x_0, $f(x)$ tiende a b. También escribimos $f(x) \to b$ cuando $x \to x_0$. (Compárese esta idea con el concepto de límite de una sucesión.) Obsérvese que si x_0 no es un punto de acumulación, no existen otros puntos de A cerca de x_0, en cuyo caso la condición es vacía.

Nota. Siempre que hagamos afirmaciones para espacios métricos, debemos preguntarnos: ¿Cuál es el enunciado correspondiente para $\mathbb{R}^n$? De la misma forma, cuando se nos dé un enunciado o un problema para $\mathbb{R}^n$, debemos preguntarnos: ¿Es esto particular de $\mathbb{R}^n$, o funciona en general en un espacio métrico, un espacio normado o un espacio vectorial?

En general, $\lim_{x\to x_0} f(x)$ *no tiene por qué existir*. Por ejemplo, sea $f: \mathbb{R} \setminus \{0\} \to \mathbb{R}$ dada por $f(x) = 1$ si $x < 0$ y $f(x) = 2$ si $x > 0$. Entonces 0 es un punto de acumulación de $\mathbb{R} \setminus \{0\}$, pero $\lim_{x\to 0} f(x)$ no existe. Sin embargo, si $f(x) = 1$ cuando $x \neq 0$ y $f(0) = 0$, entonces $\lim_{x\to 0} f(x) = 1$. Otro ejemplo es la función $f: \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) = \operatorname{sen}(1/x)$; esta función oscila cada vez más rápidamente cerca de 0 y, por lo tanto, no puede tender a límite alguno.

Si $\lim_{x\to x_0}$ *existe, entonces es único*, por lo que estamos justificados a hablar de *el* límite de f en x_0. Para demostrarlo, supongamos que $\lim_{x\to x_0} f(x) = b$ y b'. Para probar que $b = b'$, sea $\varepsilon > 0$, entonces existen $\delta_1 > 0$ y $\delta_2 > 0$ tales que $0 < d(x, x_0) < \delta_1$ implica $\rho(f(x), b) < \varepsilon/2$ y $0 < d(x, x_0) < \delta_2$ implica $\rho(f(x), b') < \varepsilon/2$. Sea $\delta = \min\{\delta_1, \delta_2\}$; entonces $0 < d(x, x_0) < \delta$ implica $\rho(b, b') \le \rho(b, f(x)) + \rho(f(x), b') < \varepsilon/2 + \varepsilon/2 = \varepsilon$, por la desigualdad triangular. Así, $\rho(b, b') < \varepsilon$ para todo $\varepsilon > 0$, luego $\rho(b, b') = 0$, lo que significa que $b = b'$. Obsérvese que en **4.1.1**, si no pedimos que x_0 sea un punto de acumulación, entonces los límites *no* tienen por qué ser únicos.

Hay algunos casos particulares de **4.1.1** en los que se usa una notación particular. Por ejemplo, supongamos que f está definida (al menos) en $A =]x_0, a] \subset \mathbb{R}$ para algún $a > x_0$ y que f tiene valores reales. Entonces

$$\lim_{x \to x_0^+} f(x) = b$$

significa el límite de f con dominio $A =]x_0, a]$. En otras palabras, para cada $\varepsilon > 0$ existe $\delta > 0$ tal que $|x - x_0| < \delta$ y $x > x_0$ implica que $|f(x) - b| < \varepsilon$. Así, estamos analizando el límite de f cuando x tiende a x_0 *por la derecha*. En forma análoga, podemos definir

$$\lim_{x \to x_0^-} f(x) = b,$$

el límite cuando x tiende a x_0 *por la izquierda*. Llamamos a estos límites, por razones obvias, *límites laterales*. El lector debería tener claro ahora cómo definir expresiones del tipo $\lim_{x \to \infty} f(x) = a$, y similares.

Ahora ya estamos en condiciones de definir la continuidad de una función en un punto.

4.1.2 Definición *Sean $A \subset M$, $f : A \to N$ y $x_0 \in A$. Decimos que f es* ***continua en*** *x_0 si x_0 no es un punto de acumulación de A o si* $\lim_{x \to x_0} f(x) = f(x_0)$.

Esta definición requiere la existencia de $\lim_{x \to x_0} f(x)$ además de especificar su valor. Puede replantearse como sigue:

f es continua en el punto x_0 de su dominio sii para todo $\varepsilon > 0$ existe $\delta > 0$ tal que para todo $x \in A$, $d(x, x_0) < \delta$ implica $\rho(f(x), (f(x_0)) < \varepsilon$.

En la definición **4.1.1** necesitamos especificar que $x \neq x_0$, pues f no tenía necesariamente que estar definida en x_0, pero en este caso no es preciso especificar $x \neq x_0$, puesto que nuestra condición claramente es válida si $x = x_0$ Obsérvese que, en el importante caso particular en que $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, *$f$ es continua en $x_0 \in A$ sii para todo $\varepsilon > 0$ existe $\delta > 0$ tal que para todo $x \in A$ que verifique $\|x - x_0\| < \delta$ tenemos que $\|f(x) - f(x_0)\| < \varepsilon$.*

4.1.3 Definición *Una función $f : A \subset M \to N$ es* ***continua en*** *el conjunto $B \subset A$ sii es continua en cada punto de B. Si sólo decimos que f es* ***continua***, *queremos dar a entender que f es continua en su dominio A.*

Hay otras formulaciones útiles del concepto de continuidad, una de las cuales aparece en el apartado **iii** del teorema siguiente y es particularmente significativa, pues sólo involucra la topología (es decir, los conjuntos abiertos), por lo que puede aplicarse a situaciones más generales.

4.1.4 Teorema *Sea $f: A \subset M \to N$ una transformación. Entonces las siguientes afirmaciones son equivalentes:*

i. *f es continua en A.*

ii. *Para cada sucesión convergente $x_k \to x_0$ en A, tenemos $f(x_k) \to f(x_0)$.*

iii. *Para cada conjunto abierto U en N, $f^{-1}(U) \subset A$ es abierto relativo a A; es decir, $f^{-1}(U) = V \cap A$ para algún conjunto abierto V.*

iv. *Para cada conjunto cerrado $F \subset N$, $f^{-1}(U) \subset A$, es cerrado relativo a A; es decir, $f^{-1}(F) = G \cap A$ para algún conjunto cerrado G.*

La condición **ii** de este teorema tiene una versión similar para límites, que se puede demostrar de la misma forma. Explícitamente, si $f: A \subset M \to N$ y x_0 es un punto de acumulación de A, entonces

$$\lim_{x \to x_0} f(x) = b \qquad \text{si y sólo si} \qquad \lim_{k \to \infty} f(x_k) = b$$

para toda sucesión $x_k \in A$ que converge a x_0, tal que $x_k \neq x_0$.

Es evidente a partir de este teorema que nuestra definición de arco continuo, dada en el capítulo 3, coincide con la continuidad como la hemos definido ahora. En §4.3 presentaremos algunos teoremas que nos permitirán establecer rápidamente la continuidad de algunas de las funciones más comunes.

Vamos ahora a analizar brevemente la plausibilidad del teorema. En primer lugar, debería estar claro que **i** y **ii** dicen lo mismo, pues **i** significa que $f(x)$ está cerca de $f(x_0)$ si x está cerca de x_0, y **ii** es igual salvo porque se permite el acercamiento de x a x_0 por medio de una sucesión. Las afirmaciones **iii** y **iv** también dicen lo mismo, si recordamos que los conjuntos abiertos son complementarios de conjuntos cerrados.

Veamos qué nos dice **iii**. Sea U un pequeño conjunto abierto que contiene a $f(x_0)$. El hecho de que $f^{-1}(U)$ sea abierto quiere decir que existe todo un disco abierto con centro en x_0 contenido en $f^{-1}(U)$. En este disco, x tiene su imagen en U, que representa puntos cercanos a $f(x_0)$. En otras palabras, si usamos U como una medida de la cercanía de $f(x)$ a $f(x_0)$, y si x está lo suficientemente cerca de x_0 (es decir, si $x \in f^{-1}(U)$), $f(x)$ estará cerca de $f(x_0)$. Por lo tanto, esto representa la misma idea expresada en **i**.

4.1.5 Ejemplo *Sea $f: \mathbb{R}^n \to \mathbb{R}^m$ la función identidad $x \mapsto x$. Muéstrese que f es continua.*

Solución Fijamos $x_0 \in \mathbb{R}^n$. Por definición, debemos encontrar $\delta > 0$ para $\varepsilon > 0$ dada tal que $\|x - x_0\| < \delta$ implique $\|f(x) - f(x_0)\| < \varepsilon$. Si elegimos $\delta = \varepsilon$, la definición se convierte en la afirmación de que $\|x - x_0\| < \varepsilon$ implica $\|x - x_0\| < \varepsilon$, lo que es una tautología. Por lo tanto, f es continua. ♦

4.1.6 Ejemplo *Sea f: $]0, \infty[\to \mathbb{R}$, $x \mapsto 1/x$. Muéstrese que f es continua.*

Solución Fijemos $x_0 \in]0, \infty[$; es decir, fijemos $x_0 > 0$. Para determinar δ, analicemos la expresión

$$|f(x) - f(x_0)| = \left|\frac{1}{x} - \frac{1}{x_0}\right| = \frac{|x_0 - x|}{|xx_0|}$$

Si $|x - x_0| < \delta$, entonces obtendríamos

$$|f(x) - f(x_0)| < \frac{\delta}{|xx_0|} = \frac{\delta}{xx_0}.$$

Si $\delta < x_0/2$, entonces $x > x_0/2$ (figura 4.1-2) de modo que $\delta/xx_0 < 2\delta/x_0^2$. Así, dado $\varepsilon > 0$, elegimos $\delta = \text{mín}(x_0/2, \varepsilon x_0^2/2)$. Entonces $f(x) - f(x_0) < \varepsilon$ si $|x - x_0| < \delta$ y, por lo tanto, f es continua en x_0. ♦

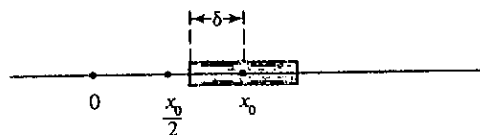


FIGURA 4.1-2 Elección de δ para el ejemplo **4.1.6**

4.1.7 Ejemplo *Sea f: $\mathbb{R}^n \to \mathbb{R}^m$ continua. Muéstrese que el conjunto $\{x \in \mathbb{R}^n \mid \|f(x)\| < 1\}$ es abierto.*

Solución El conjunto dado es igual a $f^{-1}(\{y \mid \|y\| < 1\})$, que es la imagen inversa de un conjunto abierto, por lo que el teorema **4.1.4iii** implica que el conjunto es abierto. ♦

Ejercicios de §4.1

1. **a.** Sea f: $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$. Demuéstrese que f es continua.

b. Sea f: $\mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x$. Demuéstrese que f es continua.

2. Utilícese el apartado **b** del ejercicio 1 para mostrar que si $U \subset \mathbb{R}$ es abierto, entonces $A = \{(x, y) \in \mathbb{R}^2 \,|\, x \in U\}$ es abierto.

3. Sea $f: \mathbb{R}^2 \to \mathbb{R}$ continua. Muéstrese que $A = \{(x, y) \in \mathbb{R}^2 \,|\, 0 \leq f(x, y) \leq 1\}$ es cerrado.

4. Demuéstrese en forma directa que la condición **iii** implica la condición **iv** del teorema **4.1.4**.

5. Encuéntrese un ejemplo de una función continua $f: \mathbb{R} \to \mathbb{R}$ y un conjunto abierto $U \subset \mathbb{R}$ tales que $f(U)$ *no* sea abierto.

§4.2 Imágenes de conjuntos compactos y conexos

En esta sección vamos a estudiar el comportamiento de los conjuntos compactos y conexos bajo transformaciones continuas. Es crucial distinguir entre los términos *imagen* y *preimagen* (es decir, la imagen inversa) en estos teoremas; compárense los teoremas que siguen con el teorema **4.1.1.** En esta sección, como en **§4.1**, M y N son espacios métricos con métricas d y ρ, respectivamente.

4.2.1 Teorema *Supongamos que $f: M \to N$ es continua y que $K \subset M$ es conexo. Entonces $f(K)$ es conexo. Análogamente, si K es conexo por arcos, también lo es $f(K)$.*

Intuitivamente, este teorema confirma nuestra idea de que las transformaciones continuas no pueden separar los conjuntos. Si $f(K)$ fuera disconexo, entonces f debería tener discontinuidades para separar K.

Es interesante que los conjuntos compactos cumplan el mismo tipo de condición que los conjuntos conexos (figura 4.2-1).

4.2.2 Teorema *Supongamos que $f: M \to N$ es continua y que $B \subset M$ es compacto. Entonces $f(B)$ es compacto.*

Tal vez este resultado resulte más fácil de ver si utilizamos la propiedad de Bolzano-Weierstrass. De hecho, si $f(x_k)$ es una sucesión en $f(B)$, de modo que $x_k \in B$, entonces podemos extraer una subsucesión convergente de x_k y la subsucesión correspondiente de $f(x_k)$ converge, por el teorema **4.1.4ii**.

Si $f: A \subset \mathbb{R}^n \to \mathbb{R}^m$, podemos aplicar los dos teoremas anteriores a los conjuntos $K \subset A$ y $B \subset A$, donde usamos $M = A$ como el espacio métrico del dominio y $N = \mathbb{R}^m$ como el espacio métrico del destino. Obsérvese que K es conexo como subconjunto de A sii es conexo como subconjunto de $\mathbb{R}^n$, y que lo mismo puede afirmarse para la compacidad. (El lector debe proporcionar la demostración de cada una de estas afirmaciones.)

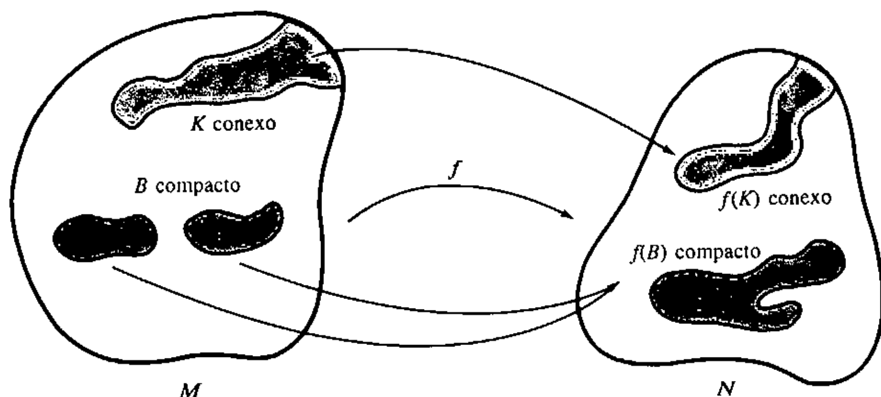


FIGURA 4.2-1 Las imágenes de conjuntos compactos (o conexos) son compactas (o conexas)

4.2.3 Ejemplo *Sea $K \subset \mathbb{R}^2$ compacto. Demuéstrese que $A = \{x \in \mathbb{R} \mid$ existe y tal que $(x, y) \in K\}$ es compacto.*

Solución Sea $f: \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x$. Entonces f es continua (véase el ejercicio 1 de §4.1). Afirmamos que $A = f(K)$, de modo que A sería compacto por el teorema **4.2.2**. Para demostrar la afirmación, sea $x \in A$. Entonces $(x, y) \in K$ para algún y, de modo que $x = f(x, y) \in f(K)$. Recíprocamente, si $x = f(x, y)$ para algún $(x, y) \in K$, entonces $x \in A$, por definición. ♦

4.2.4 Ejemplo *Encuéntrese una aplicación continua $f: \mathbb{R} \to \mathbb{R}$ y un conjunto compacto $K \subset \mathbb{R}$ tales que $f^{-1}(K)$ no sea compacto. Repítase para K conexo.*

Solución Sean $f(x) = 0$ para todo $x \in \mathbb{R}$ y $K = \{0\}$. Entonces $f^{-1}(K) = \mathbb{R}$ no es compacto. Para la segunda parte, sean $f(x) = x^2$ y $K = \{1\}$. Entonces $f^{-1}(K) = \{-1, 1\}$, que no es conexo. ♦

4.2.5 Ejemplo *Sea $f: \mathbb{R}^2 \to \mathbb{R}$ continua y sea $A = \{f(x) \mid \|x\| = 1\}$. Muéstrese que A es un intervalo cerrado, es decir, un conjunto de la forma $[a, b]$.*

Solución Claramente $A = f(S)$, donde $S = \{x \in \mathbb{R}^2 \mid \|x\| = 1\}$ es la circunferencia unidad. Ahora, S es conexo y compacto, por lo que A es conexo y compacto. Por el

ejemplo resuelto **3.3** al final del capítulo 3, A es un intervalo, pero los únicos intervalos compactos son los intervalos cerrados. ♦

Ejercicios de §4.2

1. Sea $f: \mathbb{R} \to \mathbb{R}$ continua. ¿Cuáles de los siguientes conjuntos son necesariamente cerrados, abiertos, compactos o conexos?

a. $\{x \in \mathbb{R} \mid f(x) = 0\}$

b. $\{x \in \mathbb{R} \mid f(x) > 1\}$

c. $\{f(x) \in \mathbb{R} \mid x \geq 0\}$

d. $\{f(x) \in \mathbb{R} \mid 0 \leq x \leq 1\}$

2. Verifíquense los teoremas **4.2.1** y **4.2.2** para $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 + y$, $K = B = \{(x, y) \mid x^2 + y^2 \leq 1\}$.

3. Encuéntrese un ejemplo de una aplicación continua $f: \mathbb{R} \to \mathbb{R}$ y un subconjunto cerrado $B \subset \mathbb{R}$ tales que $f(B)$ no sea cerrado. ¿Es esto posible si B es también acotado?

4. Sean A, $B \subset \mathbb{R}$ y supongamos que $A \times B \subset \mathbb{R}^2$ es conexo.

a. Demuéstrese que A es conexo.

b. Generalícese el resultado a espacios métricos.

5. Sean A, $B \subset \mathbb{R}$ y supongamos que $A \times B \subset \mathbb{R}^2$ es abierto. ¿Debe A ser abierto?

§4.3 Operaciones con transformaciones continuas

Es plausible que la composición de funciones continuas sea continua, como veremos en seguida. Recordemos que para $f: A \subset M \to N$ y $g: B \subset N \to P$, con $f(A) \subset B$, definimos la ***composición*** $g \circ f: A \to P$ como $x \mapsto g(f(x))$. Si x está cerca de x_0, entonces $(g \circ f)(x)$ está cerca de $(g \circ f)(x_0)$, pues $f(x)$ está cerca de $f(x_0)$ y por lo tanto $g(f(x))$ está cerca de $g(f(x_0))$. Véase la figura 4.3-1. Esto indica la plausibilidad del siguiente resultado.

4.3.1 Teorema *Sean M, N y P espacios métricos y supongamos que $f: A \subset M \to N$ y $g: B \subset N \to P$ son transformaciones continuas tales que $f(A) \subset B$. Entonces $g \circ f: A \subset M \to P$ es continua.*

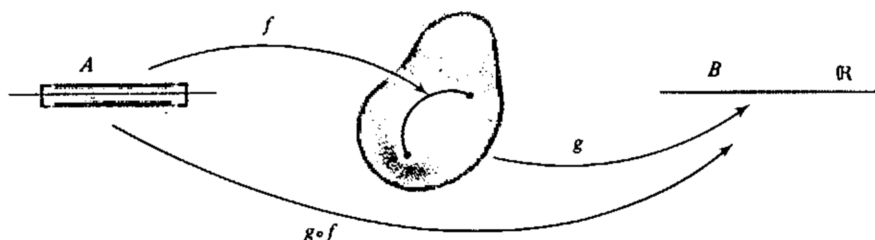


FIGURA 4.3-1 Composición de aplicaciones

Por ejemplo, la función $e^{\operatorname{sen} x}$ es continua, pues es la composición de dos funciones continuas, $f(x) = \operatorname{sen} x$ y $g(x) = e^x$.

Nota. Como ya lo hemos hecho varias veces antes, suponemos conocidas del cálculo las propiedades de las funciones elementales, tales como sen x, e^x, etcétera. Daremos un repaso de estas funciones elementales en **§5.3** y las usaremos en varios ejemplos.

El siguiente teorema da algunas propiedades fundamentales más en relación con la aritmética de los límites. En estos resultados usaremos una estructura de espacio vectorial para que tengan sentido la suma y la multiplicación por un escalar. Así, restringiremos los espacios métricos a espacios normados. Concretamente, M denotará un espacio métrico y V un espacio vectorial normado.

4.3.2 Proposición *Sean $A \subset M$ y x_0 un punto de acumulación de A.*

i. *Sean $f: A \to V$ y $g: A \to V$ funciones y definamos $f + g: A \to V$ mediante $(f+g)(x) = f(x) + g(x)$; supongamos que $\lim_{x \to x_0} f(x)$ y $\lim_{x\to x_0} g(x)$ existen y son iguales a a y b, respectivamente. Entonces $\lim_{x\to x_0} (f+g)(x)$ existe y es igual a $a + b$.*

ii. *Sean $f: A \to \mathbb{R}$ y $g: A \to V$ funciones y definamos $f \cdot g: A \to V$ mediante $(f \cdot g)(x) = f(x)g(x)$; supongamos que $\lim_{x\to x_0} f(x)$ y $\lim_{x\to x_0} g(x)$ existen y son iguales a a y b, respectivamente. Entonces $\lim_{x\to x_0} (f\cdot g)(x)$ existe y es igual a ab.*

iii. *Sean $f: A \to \mathbb{R}$ y $g: A \to V$ funciones; supongamos que $\lim_{x\to x_0} f(x)$ y $\lim_{x\to x_0} g(x)$ existen y son iguales a $a \neq 0$ y b, respectivamente. Entonces f es distinta de cero en una vecindad U de x_0, de modo que podemos definir $g/f: U \to V$ como $(g/f)(x) = g(x)/f(x)$; en ese caso $\lim_{x\to x_0} (g/f)(x)$ existe y es igual a b/a.*

Estos resultados son bastante razonables. Por ejemplo, **i** establece que si x está cerca de x_0, de modo que $f(x)$ esté cerca de a y $g(x)$ esté cerca de b, entonces $f(x) + g(x)$ estará cerca de $a + b$.

Podemos deducir algunas propiedades básicas de la aritmética de las funciones continuas.

4.3.3 Corolario *Sean $A \subset M$ y $x_0 \in A$ un punto de acumulación de A.*

i. *Sean $f : A \to V$ y $g : A \to V$ continuas en x_0; entonces la suma $f + g : A \to V$ es continua en x_0.*

ii. *Sean $f : A \to \mathbb{R}$ y $g : A \to V$ continuas en x_0; entonces el producto $f \cdot g : A \to V$ es continuo en x_0.*

iii. *Sean $f : A \to \mathbb{R}$ y $g : A \to V$ continuas en x_0,* con $f(x_0) \neq 0$; *entonces f es distinta de cero en una vecindad U de x_0 y el cociente $g/f : U \to V$ es continuo en x_0.*

Por ejemplo, hemos visto que $f(x) = x$, que transforma $\mathbb{R}$ en $\mathbb{R}$, es continua, por lo que $f(x) = x^n$ también lo es; de esto obtenemos que cualquier polinomio $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$ de $\mathbb{R}$ en $\mathbb{R}$ es continuo.

Consideremos una función de valores reales, de dos variables, $f : \mathbb{R}^2 \to \mathbb{R}$. Es crucial distinguir entre la continuidad de f (a veces llamada ***continuidad conjunta***) y la continuidad en cada variable por separado. Por ejemplo, considérese la siguiente función, que se muestra en la figura 4.3-2:

$$f(x, y) = \begin{cases} 0 & \text{si } x \neq 0 \text{ e } y \neq 0; \\ 1 & \text{si } x = 0 \text{ o } y = 0. \end{cases}$$

En cada variable por separado, f es continua en (0, 0) (las transformaciones $x \mapsto f(x, 0)$ e $y \mapsto f(0, y)$ son constantes y, por lo tanto, son continuas), pero f misma no es continua en (0, 0) (¿por qué?). En este ejemplo, obsérvese que $\lim_{(x,y)\to(0,0)} f(x, y) \neq \lim_{y\to 0}[\lim_{x\to 0} f(x, y)]$. Véanse en el ejercicio 16 al final de este capítulo algunas condiciones suficientes para que la continuidad parcial implique la continuidad conjunta.

4.3.4 Ejemplo *Sea $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x$ sen x. Muéstrese que f es continua.*

Solución Como f es el producto de dos funciones continuas, x y sen x, es continua. ♦

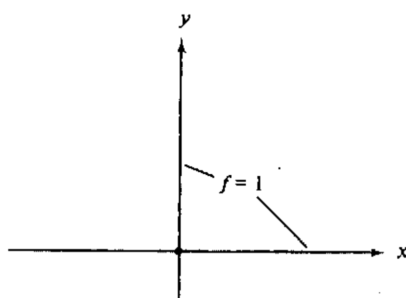


FIGURA 4.3-2 Continuidad parcial y conjunta

4.3.5 Ejemplo *Sea $f: \mathbb{R} \to \mathbb{R}^2$ continua. Muéstrese que $g(x) = f(x^2 + x^3)$ es continua.*

Solución Como g es la composición de f con la función continua $x \mapsto x^2 + x^3$, es continua. ♦

4.3.6 Ejemplo *Sea $f(x) = x^2/(1 + x)$. ¿Dónde es continua f?*

Solución La función f está definida para $x \neq -1$. Por el corolario **4.3.3iii**, f es continua en todo $x \neq -1$. ♦

Ejercicios de §4.3

1. ¿Dónde son continuas las siguientes funciones?

   a. $f(x) = x \operatorname{sen}(x^2)$.

   b. $f(x) = (x + x^2)/(x^2 - 1)$, $x^2 \neq 1$, $f(\pm 1) = 0$.

   c. $f(x) = (\operatorname{sen} x)/x$, $x \neq 0$, $f(0) = 1$.

2. Sean $a_k \to a$ y $b_k \to b$ dos sucesiones convergentes en $\mathbb{R}$. Demuéstrese que $a_k b_k \to ab$, directamente y como consecuencia de **4.33ii.**

3. Sea $A = \{x \in \mathbb{R} \mid \operatorname{sen} x = 0.56\}$. Muéstrese que A es un conjunto cerrado. ¿Es compacto?

4. Muéstrese que $f(x) = \sqrt{|x|}$ es continua.
5. Muéstrese que $f(x) = \sqrt{x^2 + 1}$ es continua.

§4.4 La acotación de las transformaciones continuas sobre conjuntos compactos

Ahora estamos listos para demostrar una importante propiedad de las funciones continuas de valores reales, llamada el teorema del "máximo-mínimo" o de "acotación". Este resultado dice que una función continua está acotada sobre un conjunto compacto y alcanza su valor máximo y su valor mínimo en algún punto del conjunto. Daremos las definiciones precisas en breve.

Para apreciar este resultado, consideremos lo que ocurre en un conjunto no compacto. En primer lugar, *una función continua no tiene por qué estar acotada*. La figura 4.4-1 muestra la función $f(x) = 1/x$ en el intervalo $]0, 1[$. Cuando x tiende a 0, la función crece sin límite, pero f sigue siendo continua (puesto que es el cociente entre 1 y la función continua $x \mapsto x$, que no se anula en $]0, 1[$).

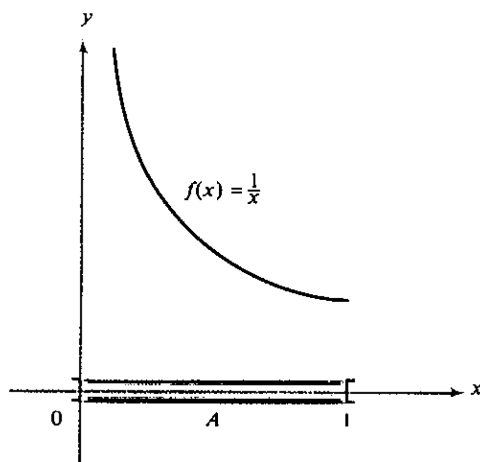


FIGURA 4.4-1 Una función continua no acotada

Además, podemos demostrar que *incluso si una función es acotada y continua, podría no alcanzar su máximo en algún punto del dominio*. La figura 4.4-2 muestra la función $f(x) = x$ en el intervalo $[0, 1[$. Esta función jamás alcanza un valor máximo, pues

aunque hay infinitos puntos tan cercanos a 1 como se quiera, no existe un punto x tal que $f(x) = 1$. Estos ejemplos deben servir para que sea plausible el hecho de que tales patologías no pueden ocurrir en una función continua sobre un conjunto compacto.

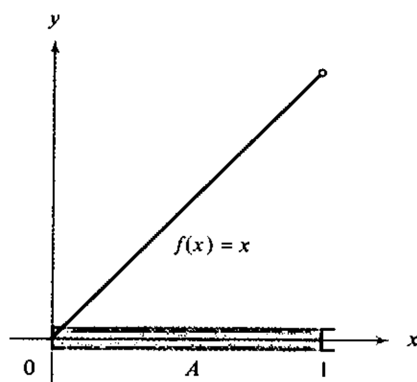


FIGURA 4.4-2 Una función sin máximo

Establezcamos ahora el teorema formalmente.

4.4.1 Teorema del máximo-mínimo *Sean (M, d) un espacio métrico, $A \subset M$ y $f: A \to \mathbb{R}$ una función continua. Sea $K \subset A$ un conjunto compacto. Entonces f está acotada en K; es decir, $B = \{f(x) \mid x \in K\} \subset \mathbb{R}$ es un conjunto acotado. Además, existen puntos $x_0, x_1 \in K$ tales que $f(x_0) = \inf(B)$ y $f(x_1) = \sup(B)$. Decimos que $\sup(B)$ es el **máximo (absoluto)** de f en K e $\inf(B)$ el **mínimo (absoluto)** de f en K.*

Obsérvese que este resultado es más profundo que los criterios usuales de derivadas para la localización de máximos y mínimos aprendidos en cálculo. Por ejemplo, existen funciones continuas en $\mathbb{R}$ que no son derivables en todo punto; tales funciones no se pueden representar mediante una curva suave y, por lo tanto, nuestra intuición no es tan clara en tales casos. (Veremos una función de este tipo en el ejercicio 20 del capítulo 5.)

Por lo general, el teorema **4.4.1** se aplica en el contexto de las funciones de varias variables; es decir, transformaciones $f: A \subset \mathbb{R}^n \to \mathbb{R}$. Sin embargo, existen algunas situaciones interesantes en las que son necesarios espacios métricos más generales (de dimensión infinita, de hecho). Una de ellas es el problema de encontrar la superficie de área mínima cuya frontera es un cierto circuito cerrado en $\mathbb{R}^3$. En este caso, M es un espacio de superficies adecuado y f asigna a cada una de ellas su área. Para minimizar f, se necesita una condición como la compacidad. En este caso, las ideas de **4.4.1** son útiles, aunque

son necesarios ciertos detalles técnicos adicionales para resolver el problema, habitualmente denominado *problema de Plateau*.

A pesar de estos aspectos técnicos, se puede aprender mucho de los experimentos físicos en los que la naturaleza realiza la minimización por nosotros. Esto se hace introduciendo alambres en soluciones jabonosas y observando las películas de jabón resultantes generadas por los alambres. La variedad de formas posibles (incluyendo varios mínimos locales) es sorprendentemente interesante. Para más información, consúltese J. Marsden y A. Tromba, *Vector Calculus*, 3a. edición, W.H. Freeman, Nueva York, 1988,[1] y S. Hildebrant y A. Tromba, *Mathematics and Optimal Form*, Scientific American Books, Nueva York, 1985.

4.4.2 Ejemplo *Proporciónese un ejemplo de una función discontinua no acotada en un conjunto compacto.*

Solución Sea $f: [0, 1] \to \mathbb{R}$ dada por $f(x) = 1/x$ si $x > 0$ y $f(0) = 0$. Está claro que esta función exhibe la misma propiedad de no acotación que la función $1/x$ en $]0, 1]$. ♦

4.4.3 Ejemplo *Verifíquese el teorema* **4.4.1** *para* $f(x) = x/(x^2 + 1)$ *en* $[0, 1]$.

Solución $f(0) = 0, f(1) = 1/2$. Verificaremos explícitamente que el máximo está en $x = 1$, y que el mínimo está en $x = 0$ (el cálculo elemental nos ayuda a determinar esto, pero daremos una verificación directa). En primer lugar, como $0 \leq x \leq 1, x/(x^2 + 1) \geq 0$, puesto que $x \geq 0$ y $x^2 + 1 \geq 1$, de modo que $f(x) \geq f(0)$ para $0 \leq x \leq 1$. Así, 0 es el mínimo. A continuación, obsérvese que $0 \leq (x - 1)^2 = x^2 - 2x + 1$, de modo que $x^2 + 1 \geq 2x$, luego para $x \neq 0$,

$$\frac{x}{x^2+1} \leq \frac{x}{2x} = \frac{1}{2}$$

por lo que $f(x) \leq f(1) = 1/2$ y entonces $x = 1$ es el punto donde se alcanza el máximo. ♦

4.4.4 Ejemplo *Muéstrese que* x_0 *y* x_1 *en el teorema* **4.4.1** *no tienen por qué ser únicos.*

Solución Sea $f(x) = 1$ para todo $x \in [0, 1]$. Entonces, cualesquiera $x_0, x_1 \in [0, 1]$ nos sirven. Por supuesto, este ejemplo es más bien trivial. Son más interesantes otros ejemplos que hemos visto en cálculo, como por ejemplo $(x^2 - 1)^2$ en $[-2, 2]$, que tiene mínimos en ± 1 y máximos en ± 2. ♦

---

[1] Edición en castellano: Marsden y Tromba, *Cálculo vectorial 3ªed.*, Addison-Wesley Iberoamericana, 1991.

Ejercicios de §4.4

1. Proporciónese un ejemplo de una función continua y acotada en todo $\mathbb{R}$ que no alcance su máximo ni su mínimo.

2. Verifíquese el teorema del máximo-mínimo para $f(x) = x^3 - x$ en $[-1, 1]$.

3. Sea $f: K \subset \mathbb{R}^n \to \mathbb{R}$ continua en un conjunto compacto K y sea $M = \{x \in K \mid f(x)$ es el máximo de f en $K\}$. Muéstrese que M es un conjunto compacto.

4. Sean $f: A \subset \mathbb{R}^n \to \mathbb{R}$ continua, $x, y \in A$ y $c: [0, 1] \to A \subset \mathbb{R}^n$ una curva que une x e y. Muéstrese que a lo largo de esta curva, f alcanza sus valores máximo y mínimo (de entre todos los valores a lo largo de la curva).

5. ¿Es válida una versión del teorema del máximo-mínimo para la función $f(x) = (\operatorname{sen} x)/x$ en $]0, \infty[$? ¿y en $[0, \infty[$?

§4.5 Teorema de los valores intermedios

El lector debe conocer el teorema de los valores intermedios del cálculo elemental; éste establece que una función continua en un intervalo alcanza todos los valores entre dos elementos dados del rango, como en la figura 4.5-1a. La función no continua de la figura 4.5-1b nunca alcanza el valor 1/2. Intuitivamente, mientras que una función discontinua puede saltar de un valor a otro, una función continua debe pasar por todos los valores intermedios. Otra forma en que puede fallar la propiedad de los valores intermedios es si el dominio A no es conexo, como lo muestra la función continua de la figura 4.5-2. Así, las hipótesis cruciales son que f sea continua y que f esté definida en una región conexa. Veremos que la demostración del teorema de valores intermedios es, de hecho, muy sencilla, debido a la forma en que hemos formulado el concepto de conexión.

4.5.1 Teorema de los valores intermedios *Sean M un espacio métrico, $A \subset M$ y $f: A \to \mathbb{R}$ continua. Supongamos que $K \subset A$ es conexo y que $x, y \in K$. Para cada número $c \in \mathbb{R}$ tal que $f(x) < c < f(y)$, existe un punto $z \in K$ tal que $f(z) = c$.*

Como los intervalos (abiertos o cerrados) son conexos, el teorema de los valores intermedios del cálculo es un caso particular de este resultado. Sin embargo, obsérvese que el teorema **4.5.1** es más general, pues se aplica, por ejemplo, a las funciones continuas de valores reales de varias variables $f(x_1, \ldots, x_n)$ definidas en todo $\mathbb{R}^n$, que es un conjunto conexo.

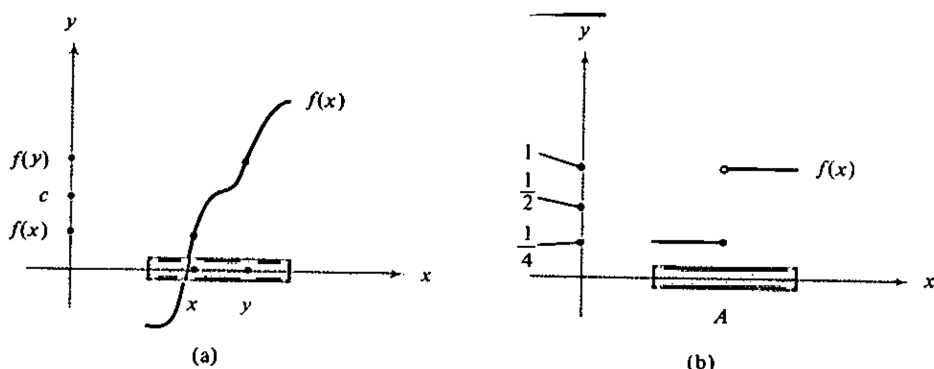


FIGURA 4.5-1 Teorema de los valores intermedios

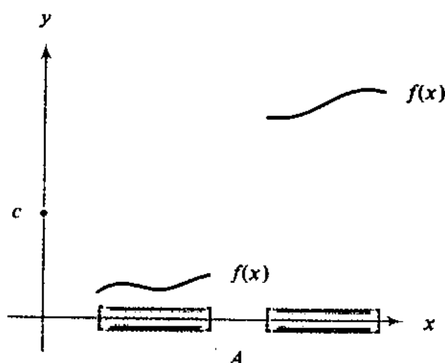


FIGURA 4.5-2 Función continua con un dominio disconexo

Puesto que los conjuntos conexos por arcos son conexos, tenemos:

4.5.2 Corolario *Sean $K \subset \mathbb{R}^n$ un conjunto conexo por arcos y $f : K \to \mathbb{R}$ continua. Sean $x, y \in K$ y $f(x) < c < f(y)$. Entonces, existe $z \in k$ tal que $f(z) = c$.*

4.5.3 Ejemplo *Demuéstrese el teorema* **4.5.1** *usando el hecho de que $f(K)$ es conexo y que los conjuntos conexos en $\mathbb{R}$ son los intervalos.*

Solución El hecho de que $f(K)$ sea conexo proviene del teorema **4.2.1**. Por lo tanto, $f(K)$ es un intervalo, tal vez no acotado. Si $f(x), f(y) \in f(K)$ son tales que $f(x) \leq f(y)$,

entonces $[f(x), f(y) \subset f(K)$ pues $f(K)$ es un intervalo. Así, si $f(x) \leq c \leq f(y)$, entonces $c \in [f(x), f(y) \subset f(K)$ y, por lo tanto, $c = f(z)$ para algún z. Daremos otra demostración en la sección de demostraciones de los teoremas al final de este capítulo. ♦

4.5.4 Ejemplo *Sea $f(x)$ un polinomio cúbico. Muéstrese que f tiene una raíz (real) x_0 (es decir, $f(x_0) = 0$).*

Solución Podemos escribir $f(x) = ax^3 - bx^2 + cx + d$, donde $a \neq 0$. Supongamos que $a > 0$. Para x grande y positivo, ax^3 es grande (y positivo) y será mayor que los demás términos, de modo que $f(x) > 0$ si x es grande. Esto requiere ciertas estimaciones precisas, pero es intuitivamente claro. Para verlo exactamente, obsérvese que

$$ax^3 + bx^2 + cx + d = ax^3\left(1 + \frac{b}{ax} + \frac{c}{ax^2} + \frac{d}{ax^3}\right)$$

y que el factor entre paréntesis tiende a 1 cuando $x \to \infty$. Análogamente, $f(x) < 0$ si x es grande y negativo. Por lo tanto, podemos aplicar el teorema de los valores intermedios con $K = \mathbb{R}$ para concluir que existe un punto x_0 tal que $f(x_0) = 0$. ♦

4.5.5 Ejemplo *Sea $f: [1, 2] \to [0, 3]$ una función continua que satisface $f(1) = 0$ y $f(2) = 3$. Muéstrese que f tiene un punto fijo. Es decir, muéstrese que existe un punto $x_0 \in [1, 2]$ tal que $f(x_0) = x_0$.*

Solución Sea $g(x) = f(x) - x$. Entonces g es continua, $g(1) = f(1) - 1 = -1$, y $g(2) = f(2) - 2 = 3 - 2 = 1$. Por lo tanto, por el teorema de los valores intermedios, g debe anularse en algún $x_0 \in [1, 2]$ y este x_0 es un punto fijo de $f(x)$. ♦

Ejercicios de §4.5

1. ¿Qué ocurre si aplicamos el método del ejemplo **4.5.4** a polinomios cuadráticos? ¿Y a los polinomios de quinto grado?

2. Sea $f: \mathbb{R}^n \to \mathbb{R}^m$ continua. Sea $\Gamma = \{(x, f(x)) \mid x \in \mathbb{R}^n\}$ la gráfica de f en $\mathbb{R}^n \times \mathbb{R}^m$. Demuéstrese que Γ es cerrada y conexa. Generalícese el resultado a espacios métricos.

3. Sea $f: [0, 1] \to [0, 1]$ continua. Demuéstrese que f tiene un punto fijo.

4. Sea $f: [a, b] \to \mathbb{R}$ continua. Muéstrese que el rango de f es un intervalo cerrado acotado.

5. Demuéstrese que no existe una función continua f, tal que $f([0, 1]) =]0, 1[$.

§4.6 Continuidad uniforme

A veces es útil tener una variante del concepto de continuidad, llamada continuidad uniforme. Esta variante es útil principalmente por razones técnicas, como el ahorro de trabajo en las demostraciones. La definición exacta es la siguiente.

4.6.1 Definición *Sean (M, d) y (N, ρ) espacios métricos, $A \subset M$, $f: A \to N$ y $B \subset A$. Decimos que f es **uniformemente continua en el conjunto B** si para cada $\varepsilon > 0$ existe $\delta > 0$ tal que $x, y \in B$ y $d(x, y) < \delta$ implican $\rho(f(x), f(y)) < \varepsilon$.*

La definición es análoga a la de continuidad, excepto que aquí necesitamos elegir δ que sirva para *todo* x, y, una vez dado ε. Para la continuidad, sólo tenemos que encontrar δ, una vez dados $\varepsilon > 0$ y un x_0 *particular*. Obviamente, si f es uniformemente continua, entonces f es continua.

Por ejemplo, considérese $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$. Entonces f es continua, pero no es uniformemente continua. De hecho, para $\varepsilon > 0$ y $x_0 > 0$ dados, el valor $\delta > 0$ que necesitamos es al menos tan pequeño como $\varepsilon/(2x_0)$ (¿por qué?), luego si x_0 es grande, δ debe ser cada vez más pequeño; es decir, no existe δ que sirva para todo x_0. Este fenómeno no puede ocurrir en conjuntos compactos, como lo muestra el siguiente teorema.

4.6.2 Teorema de la continuidad uniforme *Sean $f: A \to N$ continua y $K \subset A$ un conjunto compacto. Entonces f es uniformemente continua en K.*

El uso de conjuntos que sólo estén acotados en este teorema no sirve; consideremos, por ejemplo, la función $f(x) = 1/x$ en el conjunto no compacto $]0, 1]$. Si analizamos la demostración de que f es continua (ejemplo **4.1.6**), vemos que no es uniformemente continua. Por supuesto, no podemos extender f a una función continua en el conjunto compacto $[0, 1]$, pues f no está acotada.

4.6.3 Ejemplo *Sea $f:]0, 1] \to \mathbb{R}$ dada por $f(x) = 1/x$. Muéstrese que f es uniformemente continua en $[a, 1]$ para $a > 0$.*

Solución Esto se sigue del teorema de continuidad uniforme, pues $[a, 1]$ es un conjunto compacto y f es continua en $]0, 1]$ y, por lo tanto, también en $[a, 1]$. ♦

Aunque en la siguiente sección daremos un repaso del cálculo diferencial, vamos a usar algunas de sus ideas en este momento, pues son necesarias para el siguiente resultado.

4.6.4 Ejemplo *Sea f: $]a, b[\to \mathbb{R}$ derivable y supongamos que existe una constante $M > 0$ tal que $|f'(x)| \leq M$ para todo x que satisfaga $a < x < b$. Aquí, a o b pueden ser $\pm\infty$ y f' representa la derivada de f. Muéstrese que f es uniformemente continua en $]a, b[$.*

Solución La definición de continuidad uniforme requiere estimar la diferencia $|f(x) - f(y)|$ en términos de $|x - y|$ y esto sugiere usar el teorema del valor medio. En efecto,

$$f(x) - f(y) = f'(x_0)(x - y)$$

para algún x_0 entre x e y. Por lo tanto,

$$|f(x) - f(y)| \leq M|x - y|;$$

una transformación con esta propiedad se denomina ***Lipschitz*** (encontraremos todavía varias veces esta propiedad). Dado $\varepsilon > 0$, sea $\delta = \varepsilon/M$. Entonces, $|x - y| < \delta$ implica

$$|f(x) - f(y)| < M \cdot \delta = \frac{M \cdot \varepsilon}{M} = \varepsilon.$$

En consecuencia, f es uniformemente continua. ♦

La intuición sugerida por este resultado puede arrojar cierta luz sobre la continuidad uniforme. Específicamente, este resultado dice que si la pendiente de la gráfica de una función está acotada, entonces ésta es uniformemente continua. Las pendientes no acotadas sugieren que uno debería esperar que falle la continuidad uniforme. Esto sirve con frecuencia como una buena guía al analizar funciones específicas o sus gráficas. Sin embargo, como muestra el ejercicio 7, tal guía no es infalible.

4.6.5 Ejemplo *Muéstrese que* sen $x : \mathbb{R} \to \mathbb{R}$ *es uniformemente continua.*

Solución $d(\text{sen } x)/dx = \cos x$ está acotado en valor absoluto por 1, luego por **4.6.4**, sen x es uniformemente continua. ♦

Ejercicios de §4.6

1. Demuéstrese la conclusión del ejemplo **4.6.3** directamente a partir de la definición.

2. Demuéstrese que $f(x) = 1/x$ es uniformemente continua en $[a, \infty[$ para $a > 0$.

3. ¿Debe ser uniformemente continua una función continua y acotada en $\mathbb{R}$?

4. Si f y g son aplicaciones uniformemente continuas de $\mathbb{R}$ a $\mathbb{R}$, ¿debe la función producto $f \cdot g$ ser uniformemente continua? ¿Qué ocurre si f y g son acotadas?

5. Sea $f(x) = |x|$. Muéstrese que $f: \mathbb{R} \to \mathbb{R}$ es uniformemente continua.

6. **a.** Muéstrese que $f: \mathbb{R} \to \mathbb{R}$ *no* es uniformemente continua sii existen $\varepsilon > 0$ y sucesiones x_n e y_n tales que $|x_n - y_n| < 1/n$ y $|f(x_n) - f(y_n)| \geq \varepsilon$. Generalícese esta afirmación a espacios métricos.

b. Use **a** en $\mathbb{R}$ para demostrar que $f(x) = x^2$ no es uniformemente continua.

7. Sea $f(x) = \sqrt{x}$.

a. Muéstrese que f es uniformemente continua en el intervalo $[0, 1]$.

b. Analícese la relación entre la continuidad uniforme y las pendientes acotadas, a la vista de este ejemplo.

§4.7 Derivación de funciones de una variable

Supondremos que el lector está familiarizado con las ideas generales del cálculo de funciones de una variable. Nuestro propósito aquí es recordar y precisar algunos de los puntos teóricos básicos. Incluimos algunas de las demostraciones en el texto, ya que muchas de ellas serán repasos.

4.7.1 Definición *Sea f una función definida en un intervalo abierto que contiene a $x_0 \in \mathbb{R}$. Decimos que f es **derivable en** x_0 si*

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

*existe. Llamamos a $f'(x_0)$ la **derivada** de f en x_0.*

Reescribiendo esta condición como

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - f'(x_0) \cdot (x - x_0)}{x - x_0} = 0,$$

puede verse que la recta $y = f(x_0) + f'(x_0)(x - x_0)$, llamada la ***recta tangente a la gráfica de f en*** x_0, es una buena aproximación a f cerca de x_0, reescribiéndola como

$$\lim_{\Delta x \to 0} \left[\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} - f'(x_0)\right] = 0,$$

vemos que $f'(x_0)$, siendo el límite de las pendientes de rectas secantes, se puede interpretar como la pendiente de la recta tangente a la gráfica de f en $(x_0, f(x_0))$, como muestra la figura 4.7-1.

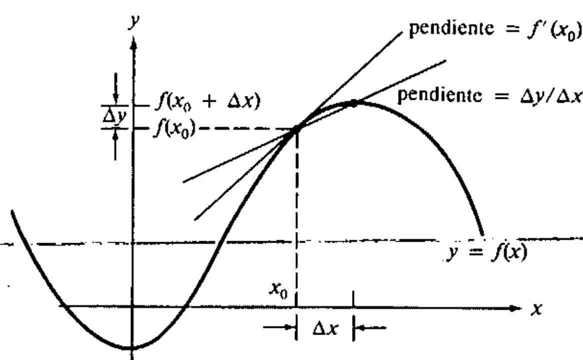


FIGURA 4.7-1 $f'(x_0)$ es la pendiente de la recta tangente

Otra forma de escribir la definición que evita la división por Δx (y la exclusión $\Delta x \neq 0$) es la siguiente: para cada $\varepsilon > 0$ existe $\delta > 0$ tal que $|\Delta x| < \delta$ implica

$$|f(x_0 + \Delta x) - f(x_0) - f'(x_0)\Delta x| \leq \varepsilon|\Delta x|.$$

Supongamos que el lector está familiarizado con la notación de Leibniz dy/dx para la derivada. A continuación damos la relación entre la derivabilidad y la continuidad:

4.7.2 Proposición *Si f es derivable en x_0, entonces f es continua en x_0.*

Demostración Tenemos que

$$\begin{aligned}\lim_{x \to x_0} f(x) &= \lim_{x \to x_0} \left[\frac{f(x) - f(x_0)}{x - x_0}(x - x_0) + f(x_0)\right] \\ &= f'(x_0) \cdot 0 + f(x_0) = f(x_0). \quad \blacksquare\end{aligned}$$

A continuación damos otra demostración de esta proposición, que necesitaremos más adelante. Dado $\varepsilon > 0$ existe $\delta > 0$ tal que

$$|f(x_0 + \Delta x) - f(x_0) - f'(x_0)\Delta x| \leq \varepsilon|\Delta x|$$

si $|\Delta x| < \delta$. Así, por la desigualdad triangular,

$$|f(x_0 + \Delta x) - f(x_0)| \leq |f'(x_0)\Delta x| + \varepsilon|\Delta x|.$$

Si elegimos $\varepsilon = 1$,

$$|f(x_0 + \Delta x) - f(x_0)| \leq (f'(x_0) + 1)|\Delta x|$$

si $|\Delta x| < \delta$. Esto demuestra algo más que la continuidad: prueba que f tiene la propiedad ***de Lipschitz*** en x_0, es decir, que existen constantes $M \geq 0$ y $c > 0$ tales que

$$|f(x_0 + \Delta x) - f(x_0)| \leq M|\Delta x|$$

si $|\Delta x| < c$. La continuidad es consecuencia de esto (dado $\varepsilon > 0$, tómese $\delta = \varepsilon/M$). Esta desigualdad implica que la gráfica está *dentro* de la región de la "pajarita"[2] de la figura 4.7-2.

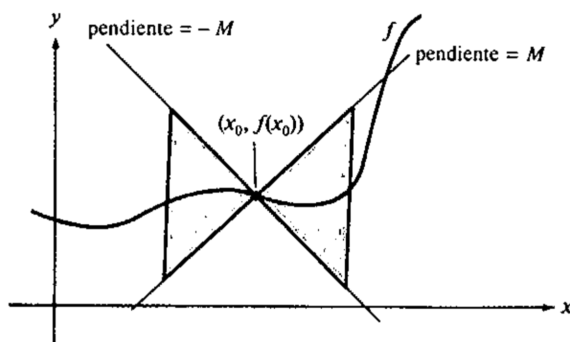


FIGURA 4.7-2 Significado geométrico de la propiedad de Lipschitz

4.7.3 Ejemplo *Sea $f(x) = |x|$. Muéstrese que f es continua, pero no derivable, en $x_0 = 0$.*

Solución El límite del cociente de diferencias,

$$\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \lim_{\Delta x \to 0} \frac{|\Delta x|}{\Delta x},$$

[2] O corbatín. (*N. del T.*)

no existe, pues $|\Delta x|/\Delta x$ es 1 si $\Delta x > 0$ y -1 si $\Delta x < 0$. La función es claramente continua en $x_0 = 0$ (sea $\varepsilon = \delta$ en la definición de continuidad). ♦

4.7.4 Ejemplo *Sea $f(x) = x^2$. Calcúlese $f'(x_0)$ a partir de la definición.*

Solución Tenemos que

$$\lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{(x + \Delta x)^2 - x^2}{\Delta x} = \lim_{\Delta x \to 0} [2x + \Delta x] = 2x,$$

y, por lo tanto, $f'(x) = 2x$; así, f es derivable en todo $x \in \mathbb{R}$. ♦

Por supuesto, en los problemas prácticos no calculamos las derivadas mediante la definición, sino que utilizamos las reglas usuales del cálculo:

4.7.5 Teorema *Supongamos que f y g son derivables en x_0 y que $k \in \mathbb{R}$. Entonces, kg, $f + g$, fg y f/g (si $g(x_0) \neq 0$) son derivables en x_0 y*

i. $(kf)'(x_0) = k(f'(x_0))$ ***regla de la multiplicación por una constante***

ii. $(f + g)'(x_0) = f'(x_0) + g'(x_0)$ ***regla de la suma***

iii. $(fg)'(x_0) = f(x_0)g'(x_0) + f'(x_0)g(x_0)$ ***regla del producto***

iv. $\left(\dfrac{f}{g}\right)'(x_0) = \dfrac{g(x_0)f'(x_0) - f(x_0)g'(x_0)}{[g(x_0)]^2}$ ***regla del cociente***

La demostración de cada caso es una inmediata manipulación de límites, que deberá escribir el lector. La siguiente regla requiere un poco más de ingenio:

4.7.6 Regla de la cadena *Si f es derivable en x_0 y g es derivable en $f(x_0)$, entonces $g \circ f$ es derivable en x_0 y*

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Esta demostración aparece al final del capítulo. Por supuesto, por medio de estas reglas (**4.7.2**, **4.7.5** y **4.7.6**) es como aprendemos a derivar, por ejemplo, las funciones racionales.

Si f es derivable en $]a, b[$ y f' es continua, decimos que f es de ***clase*** C^1. Por supuesto, podemos derivar f' de nuevo, suponiendo que sea derivable, para obtener la segunda derivada, f''. Si f'' es continua, decimos que f es de ***clase*** C^2, etcétera.

4.7.7 Definición *Decimos que una función f definida en una vecindad de x_0 es* ***creciente*** *(respectivamente,* ***estrictamente creciente****) en x_0 si existe un intervalo $]a, b[$ que contiene a x_0 tal que:*

i. *Si $a < x < x_0$, entonces $f(x) \leq f(x_0)$; respectivamente, $f(x) < f(x_0)$.*

ii. *Si $x_0 < x < b$, entonces $f(x) \geq f(x_0)$; respectivamente, $f(x) > f(x_0)$.*

De forma similar, f es ***decreciente*** *(respectivamente,* ***estrictamente decreciente****) en x_0 si existe un intervalo $]a, b[$ que contiene a x_0 tal que:*

i. *Si $a < x < x_0$, entonces $f(x) \geq f(x_0)$; respectivamente, $f(x) > f(x_0)$.*

ii. *Si $x_0 < x < b$, entonces $f(x) \leq f(x_0)$; respectivamente, $f(x) < f(x_0)$.*

4.7.8 Teorema *Supongamos que f es derivable en x_0.*

i. *Si f es creciente en x_0, entonces $f'(x_0) \geq 0$.*

ii. *Si f es decreciente en x_0, entonces $f'(x_0) \leq 0$.*

iii. *Si $f'(x_0) > 0$, entonces f es estrictamente creciente en x_0.*

iv. *Si $f'(x_0) < 0$, entonces f es estrictamente decreciente en x_0.*

Posponemos la demostración para el final del capítulo. Con este resultado, podemos relacionar el cálculo diferencial con los máximos y los mínimos.

4.7.9 Proposición *Si $f:]a, b[\to \mathbb{R}$ es derivable en $c \in]a, b[$ y f tiene un máximo (o un mínimo) en c, entonces $f'(c) = 0$.*

Demostración Sea f una función con un máximo en c. Entonces, para $h > 0$, $[f(c + h) - f(c)]/h \leq 0$, así que, haciendo $h \to 0$, $h > 0$, obtenemos $f'(c) \leq 0$. De forma análoga, para $h < 0$ obtenemos $f'(c) \geq 0$, por lo que $f'(c) = 0$. ■

A continuación reunimos la proposición anterior con el teorema del máximo-mínimo de **§4.4**.

4.7.10 Teorema de Rolle *Si $f: [a, b] \to \mathbb{R}$ es continua, f es derivable en $]a, b[$ y $f(b) = f(a) = 0$, entonces existe un número $c \in]a, b[$ tal que $f'(c) = 0$.*

Demostración Si $f(x) = 0$ para todo $x \in [a, b]$, podemos elegir cualquier c. Por lo tanto, supondremos que f no es idénticamente nula. Por el teorema del máximo-mínimo (**4.4.1**), sabemos que existe un punto c_1 donde f alcanza su máximo y un punto c_2 donde f alcanza su mínimo. Como f no es idénticamente nula y $f(a) = f(b) = 0$, al menos uno de los puntos c_1, c_2 está en $]a, b[$. Si $c_1 \in]a, b[$, obtenemos $f'(c_1) = 0$, por la proposición **4.7.9**; análogamente para c_2. ■

4.7.11 Teorema del valor medio *Si $f: [a, b] \to \mathbb{R}$ es continua y derivable en $]a, b[$, existe un punto $c \in]a, b[$ tal que $f(b) - f(a) = f'(c)(b - a)$.*

Demostración Se define $\varphi(x) = f(x) - f(a) - (x - a)[f(b) - f(a)]/(b - a)$ (véase la figura 4.7-3) y se aplica el teorema de Rolle. ■

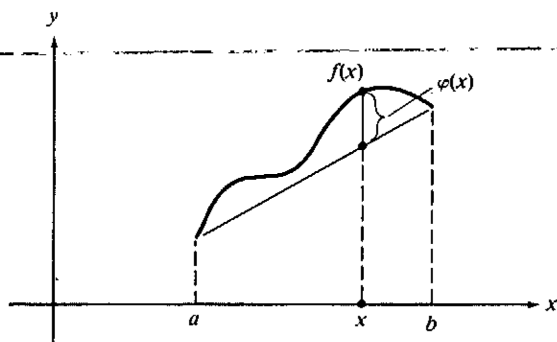


FIGURA 4.7-3 Teorema del valor medio

4.7.12 Corolario *Si $f: [a, b] \to \mathbb{R}$ es continua y $f' = 0$ en $]a, b[$, entonces f es constante.*

Demostración Aplicamos el teorema del valor medio a f en $[a, x]$ para obtener un punto c tal que $f(x) - f(a) = f'(c)(x - a) = 0$, de modo que $f(x) = f(a)$ para todo $x \in [a, b]$ y, en consecuencia, f es constante. ■

4.7.13 Ejemplo *Sea $f:]a, b[\to \mathbb{R}$ derivable y $|f'(x)| \leq M$. Demuéstrese que $|f(x) - f(y)| \leq M\,|x - y|$ para todo $x, y \in]a, b[$.*

Solución Por el teorema del valor medio,

$$f(x) - f(y) = f'(c)(x - y)$$

para algún $c \in]x, y[$. Obtenemos el resultado al tomar valores absolutos en esta expresión. ♦

Para ver la importancia del teorema del valor medio, consideremos una función $f: [a, b] \to \mathbb{R}$ derivable en $]a, b[$ y continua en $[a, b]$, y supongamos que $f'(x) \geq 0$ para todo $x \in]a, b[$. Entonces, dados $x_1, x_2 \in]a, b[$, tales que $x_1 < x_2$, podemos aplicar el teorema del valor medio a f restringida a $[x_1; x_2]$. Vemos que existe $c \in]x_1, x_2[$ tal que $f(x_2) - f(x_1) = f'(c)(x_2 - x_1)$. Pero $f'(c) \geq 0$, por hipótesis, luego $f(x_2) \geq f(x_1)$. Hemos probado, pues, que si $x_1, x_2 \in [a, b]$ y $x_1 < x_2$, entonces $f(x_1) \leq f(x_2)$; es decir, f es creciente en $[a, b]$. Esto demuestra el apartado **i** de la siguiente proposición.

4.7.14 Proposición *Supongamos que f es continua en* $[a, b]$ *y derivable en* $]a, b[$.

i. *Si* $f'(x) \geq 0$ *para todo* $x \in]a, b[$, *entonces f es creciente en* $[a, b]$.

ii. *Si* $f'(x) \leq 0$ *para todo* $x \in]a, b[$, *entonces f es decreciente en* $[a, b]$.

iii. *Si* $f'(x) > 0$ *para todo* $x \in]a, b[$, *entonces f es estrictamente creciente en* $[a, b]$.

iv. *Si* $f'(x) < 0$ *para todo* $x \in]a, b[$, *entonces f es estrictamente decreciente en* $[a, b]$.

v. *Si* $f'(x) = 0$ *para todo* $x \in]a, b[$, *entonces* $f(x) = k$ *(k es una constante) para todo* $[a, b]$.

Ahora, supongamos que $f:]a, b[\to \mathbb{R}$ es tal que $f'(x) > 0$ para todo $x \in]a, b[$ o que $f'(x) < 0$ para todo $x \in]a, b[$. Entonces f es estrictamente creciente o estrictamente decreciente en $]a, b[$, de modo que f tiene una función inversa. El siguiente teorema nos permite calcular la derivada de f^{-1} en términos de la derivada de f.

4.7.15 Teorema de la función inversa *Supongamos que* $f:]a, b[\to \mathbb{R}$ *y que, o bien* $f'(x) > 0$ *para todo* $x \in]a, b[$, *o bien* $f'(x) < 0$ *para todo* $x \in]a, b[$. *Entonces* $f:]a, b[\to \mathbb{R}$ *es una biyección sobre su recorrido,* f^{-1} *es derivable en su dominio y* $(f^{-1})'(y) = 1/f'(x)$, *donde* $f(x) = y$.

La proposición **4.7.14** también es útil para el estudio de máximos y mínimos; por ejemplo:

4.7.16 Proposición *Supongamos que f es continua en* $[a, b]$, *que es derivable dos veces en* $]a, b[$, *y que* $x_0 \in\ ]a, b[$.

i. *Si* $f'(x_0) = 0$ *y* $f''(x_0) > 0$, *entonces* x_0 *es un mínimo local estricto de f.*

ii. *Si* $f'(x_0) = 0$ *y* $f''(x_0) < 0$, *entonces* x_0 *es un máximo local estricto de f.*

Lo que ocurre en **i** es que $f''(x) > 0$ implica que f' es estrictamente creciente en x_0, por lo que $f' < 0$ justo a la izquierda de x_0 y $f' > 0$ justo a la derecha. Después usamos **4.7.14** para obtener el resultado.

Ejercicios de §4.7

1. Proporciónese un ejemplo de una función definida en $\mathbb{R}$ que sea continua en todos los puntos y que no sea derivable en exactamente n puntos dados $x_1, \ldots, x_n$.

2. ¿Se puede aplicar el teorema del valor medio a $f(x) = \sqrt{x}$ en $[0, 1]$? ¿Y a $g(x) = \sqrt{|x|}$ en $[-1, 1]$?

3. Sea f un polinomio no constante tal que $f(0) = f(1)$. Demuéstrese que f tiene un punto de mínimo o máximo local en el intervalo abierto $]0, 1[$.

4. Un cubo de caucho de material incompresible es empujado en todas las caras con una fuerza T. El material se alarga por un factor ν en dos direcciones y se contrae por un factor ν^{-2} en la otra dirección. Mediante un balance de fuerzas, se puede obtener la ***ecuación de Rivlin***:

$$\nu^3 - \frac{T\nu^2}{2\alpha} + 1 = 0,$$

donde a es una constante estrictamente positiva (análoga a la constante de recuperación en un resorte). Muéstrese que la ecuación de Rivlin tiene una solución (real) si $T < 6\sqrt{2}\alpha$ y tres soluciones si $T > 6\sqrt{2}\alpha$.

5. Sea f continua en $[3, 5]$ y derivable en $]3, 5[$, tal que $f(3) = 6$ y $f(5) = 10$. Demuéstrese que existe un punto x_0 en el intervalo abierto $]3, 5[$ tal que la recta tangente a la gráfica de f en x_0 pasa por el origen. Ilústrese el resultado mediante un dibujo.

§4.8 Integración de funciones de una variable

Sean $A \subset \mathbb{R}$ un conjunto acotado y $f: A \to \mathbb{R}$ una función acotada. Cuando decimos que queremos *integrar* la función f en el conjunto A, esto significa que queremos *determinar el área* por debajo de la gráfica de f (véase la figura 4.8-1). Si $f \geq 0$, ésta es el *área literal*; en general, es el *área con signo*. Para ello, obsérvese primero que, como A está acotado, existe un intervalo cerrado $[a, b] \supset A$. Considéresemos que f está definida en todo el intervalo $[a, b]$, haciendo que f sea igual a cero en $[a, b] \backslash A$. A continuación, hagamos una partición de $[a, b]$, lo que significa elegir un entero n y un conjunto de puntos $x_0 = a, x_1, \ldots, x_{n-1}, x_n = b$ de tal modo que $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$. Denotemos esa partición P; es decir, $P = \{x_0, \ldots, x_n\}$. Formemos entonces las dos sumas

$$U(f, P) = \sum_{i=0}^{n-1} (\sup\{f(x) \mid x \in [x_i, x_{i+1}]\})(x_{i+1} - x_i)$$

y

$$L(f, P) = \sum_{i=0}^{n-1} (\inf\{f(x) \mid x \in [x_i, x_{i+1}]\})(x_{i+1} - x_i),$$

llamadas ***sumas superior*** e ***inferior***, respectivamente. La primera es la suma sobre todos los intervalos $[x_i, x_{i+1}]$ del supremo de f en cada intervalo, por la longitud de dicho intervalo y su valor es igual al área de la región sombreada que se muestra en la figura 4.8-1. Como hemos supuesto que f está acotada, el supremo existe en cada intervalo. La segunda suma es la suma sobre todos los intervalos $[x_i, x_{i+1}]$ del ínfimo de f en cada intervalo, por la longitud de dicho intervalo, y es la región que se muestra en la figura 4.8-1. El hecho de que f esté acotada garantiza de nuevo que el ínfimo existe.

Como f está acotada (digamos que $-M \leq f \leq M$), vemos que

$$-(b - a)M \leq L(f, P) \leq U(f, P) \leq (b - a)M$$

para cualquier partición P de $[a, b]$. Sean

$$S = \inf\{U(f, P) \mid P \text{ es cualquier partición}\}$$

y

$$s = \sup\{L(f, P) \mid P \text{ es cualquier partición}\}.$$

Si observamos de nuevo la figura 4.8-1, parece razonable esperar que, cuando el tamaño de los intervalos de P sea cada vez más pequeño, $U(f, P)$ decrezca, mientras que $L(f, P)$ crezca. En el límite del tamaño decreciente de los intervalos de P, los números $U(f, P)$ y $L(f, P)$ deben converger a un valor común. Esto nos lleva a la siguiente definición.

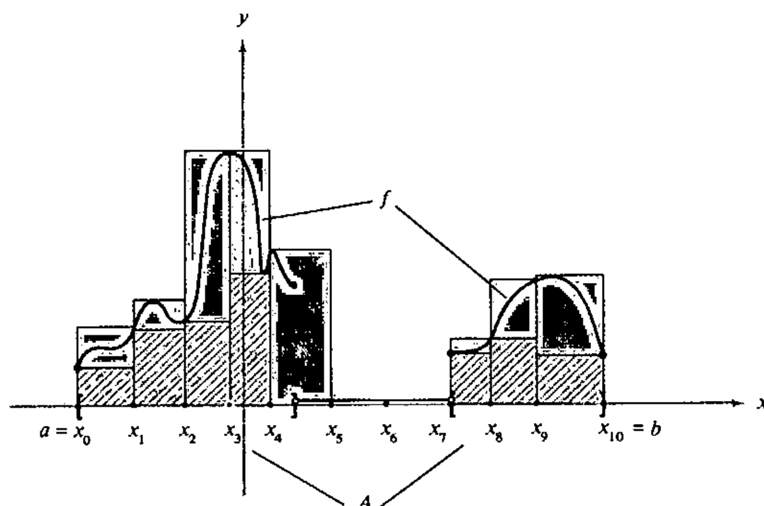


FIGURA 4.8-1 Las sumas superiores e inferiores de esta función

4.8.1 Definición *Decimos que f es* ***integrable Riemann*** *(o solamente* ***integrable****, o que la* ***integral existe****) si $s = S$. El valor común $s = S$ se denota $\int_A f$ o $\int_A f(x)dx$. Si $A = [a, b]$, escribimos*

$$\int_a^b f(x)\,dx = \int_A f = \int_a^b f.$$

Obsérvese que la integrabilidad no implica realmente propiedades de derivabilidad o continuidad de f. De hecho, algunas funciones exageradamente discontinuas pueden ser integrables.

Es probable que una parte del procedimiento intrigue al lector. ¿Por qué insistimos en que $\inf\{U(f, P)\} = \sup\{L(f, P)\}$? En principio, podríamos creer que esta relación siempre se cumple. Sin embargo, esto no siempre es así, como lo muestra el siguiente ejemplo.

4.8.2 Ejemplo *Calcúlese* $\inf\{U(f, P)\}$ *y* $\sup\{L(f, P)\}$ *para la función* $f: [0, 1] \subset \mathbb{R} \to \mathbb{R}$ *dada por* $f(x) = 1$ *si x es irracional y* $f(x) = 0$ *si x es racional.*

Solución Afirmamos que $\inf\{U(f, P)\} = 1$ y que $\sup\{L(f, P)\} = 0$. De hecho, en cualquier intervalo, f siempre vale 1 en algunos puntos y cero en otros, por lo que el ínfimo en cada intervalo es cero y el supremo es 1. Por lo tanto, la integral de esta función sobre el conjunto [0, 1] no existe para nuestros fines. En un contexto más avanzado, se

puede definir la integral de una función tan "patológica", pero nosotros trabajaremos principalmente con funciones "decentes", para las cuales existe la integral. ♦

4.8.3 Ejemplo *Supongamos que* $f: [a, b] \to \mathbb{R}$ *es integrable (Riemann) y que* $f \geq 0$. *Muéstrese que* $\int_a^b f(x)dx \geq 0$.

Solución Por definición, la integral es el ínfimo de las sumas de la forma

$$\sum_{i=0}^{n-1} \left(\sup_{x \in [x_i, x_{i+1}])} f(x) \right) \cdot (x_{i+1} - x_i)$$

sobre todas las particiones. Pero cada una de estas sumas $U(f, P)$ es no negativa, pues $f \geq 0$. Por lo tanto, la integral es ≥ 0, pues el ínfimo de un conjunto de números no negativo también es no negativo. ♦

El primer resultado es un teorema que nos da una pista sobre qué funciones son integrables. En el capítulo 8 estudiaremos este problema con más detalle.

4.8.4 Teorema

i. *Si* $f: [a, b] \to \mathbb{R}$ *está acotada y es continua en todos los puntos de* $[a, b]$ *excepto en un número finito de ellos, entonces es integrable en* $[a, b]$.

ii. *Cualquier función creciente o decreciente en* $[a, b]$ *es integrable en* $[a, b]$.

También podemos demostrar la siguiente proposición, que nos da algunas reglas de integración.

4.8.5 Proposición

i. *Si* f *es integrable en* $[a, b]$ *y* $k \in \mathbb{R}$, *entonces* kf *es integrable en* $[a, b]$ *y* $\int_a^b kf = k \int_a^b f$.

ii. *Si* f *y* g *son integrables en* $[a, b]$, *entonces* $f + g$ *es integrable en* $[a, b]$ *y* $\int_a^b (f + g) = \int_a^b f + \int_a^b g$.

iii. *Si* f *y* g *son integrables en* $[a, b]$ *y* $f(x) \leq g(x)$ *para todo* $x \in [a, b]$, *entoces* $\int_a^b f \leq \int_a^b g$.

iv. *Si* f *es integrable en* $[a, b]$ *y en* $[b, c]$, *entonces* f *es integrable en* $[a, c]$ *y* $\int_a^c f = \int_a^b f + \int_b^c f$.

Cada una de estas propiedades es plausible, en términos de la interpretación de la integral como el área con signo debajo de la gráfica. He aquí una consecuencia útil de **iii.** Obsérvese que $-|f| \le f \le |f|$, de modo que $-\int_a^b |f(x)| dx \le \int_a^b f(x) dx \le \int_a^b |f(x)| dx$. Así pues,

$$\left| \int_a^b f(x)\, dx \right| \le \int_a^b |f(x)|\, dx.$$

Vamos a introducir la siguiente notación para usarla en el desarrollo de las demostraciones y en los ejemplos:

$$\overline{\int_a^b} f(x)\, dx = \inf\{U(f, P) \mid P \text{ es cualquier partición}\}, \quad \textbf{\textit{la integral superior}}$$

y

$$\underline{\int_a^b} f(x)\, dx = \sup\{L(f, P) \mid P \text{ es cualquier partición}\}, \quad \textbf{\textit{la integral inferior}}$$

de modo que la definición de integrabilidad se lee

$$\underline{\int_a^b} f(x)\, dx = \overline{\int_a^b} f(x)\, dx.$$

En la sección de demostraciones de los teoremas mostraremos que

$$\underline{\int_a^b} f(x)\, dx \le \overline{\int_a^b} f(x)\, dx$$

siempre se verifica; así, *la integrabilidad afirma que la desigualdad estricta no es cierta.*

Vamos ahora a calcular algunas integrales directamente de la definición.

4.8.6 Ejemplo *Calcúlese* $\int_1^3 1 dx$.

Solución Consideremos una partición $P: 1 = x_0 < x_1 < \cdots < x_n = 3$. Como el ínfimo de 1 sobre cada subintervalo es 1,

$$L(1, P) = \sum_{i=1}^{n} 1 \cdot (x_i - x_{i-1}) = x_n - x_0 = 3 - 1 = 2.$$

De la misma forma, cada supremo sobre $[x_i, x_{i+1}]$ es 1, por lo que $U(1, P) = 2$. Así,

$$\underline{\int_1^3} 1\, dx = 2, \quad \overline{\int_1^3} 1\, dx = 2, \quad \text{y entonces} \quad \int_1^3 1\, dx = 2. \quad \blacklozenge$$

4.8.7 Ejemplo *Calcúlese* $\int_0^1 x\, dx$.

Solución Definamos P_n como $x_i = i/n$ y sea m_i el ínfimo de $f(x)$ sobre $[x_{i-1}, x_i]$. En este caso, $m_i = (i-1)/n$ y entonces

$$L(x, P_n) = \sum_{i=1}^{n} \left(\frac{i-1}{n}\right) \cdot \left(\frac{i}{n} - \frac{i-1}{n}\right) = \sum_{i=1}^{n} \frac{i-1}{n} \cdot \frac{1}{n}$$
$$= \frac{1}{n^2} \sum_{i=1}^{n} (i-1) = \frac{1}{n^2} \cdot \frac{n(n-1)}{2}$$

(pues $\sum_{i=1}^{n} i = n(n+1)/2$). Así,

$$\lim_{n\to\infty} L(x, P_n) = \lim_{n\to\infty} \frac{n-1}{2n} = \frac{1}{2}.$$

Por lo tanto, $\underline{\int_0^1} x\, dx \geq 1/2$, ya que $\underline{\int_0^1} x\, dx$ es una cota superior de las sumas inferiores. Análogamente, el supremo de $f(x)$ sobre $[x_{i-1}, x_i]$ es $M_i = i/n$, de donde

$$U(x, P_n) = \sum_{i=1}^{n} \left(\frac{i}{n}\right) \cdot \left(\frac{i}{n} - \frac{i-1}{n}\right) = \sum_{i=1}^{n} \frac{i}{n} \cdot \frac{1}{n}$$
$$= \frac{1}{n^2} \sum_{i=1}^{n} i = \frac{1}{n^2} \cdot \frac{n(n+1)}{2}.$$

Así, $\lim_{n\to\infty} U(x, P_n) = 1/2$ y entonces $\overline{\int_0^1} x\, dx \leq 1/2$. En consecuencia,

$$\frac{1}{2} \leq \underline{\int_0^1} x\, dx \leq \overline{\int_0^1} x\, dx \leq \frac{1}{2}$$

y por lo tanto

$$\underline{\int_0^1} x\, dx = \overline{\int_0^1} x\, dx = \frac{1}{2} = \int_0^1 x\, dx. \quad \blacklozenge$$

Ahora que hemos caracterizado una gran clase de funciones integrables, seguiremos buscando una forma *práctica* de calcular integrales. Por supuesto, la respuesta en una dimensión es el uso de primitivas y las técnicas usuales de integración. Las técnicas para dimensiones superiores están en los capítulos 8 y 9.

Para $f : [a, b] \to \mathbb{R}$, una ***primitiva*** de f es una función continua $F : [a, b] \to \mathbb{R}$ tal que F es derivable en $]a, b[$ y $F'(x) = f(x)$ para todo $a < x < b$. El siguiente teorema proporciona un método eficaz para el cálculo de la integral de una clase amplia de funciones.

4.8.8 Teorema fundamental del cálculo *Sea $f: [a, b] \to \mathbb{R}$ continua. Entonces f tiene una primitiva F y*

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Si G es otra primitiva de f, también tenemos que $\int_a^b f(x)\,dx = G(b) - G(a)$.

4.8.9 Ejemplo $\int_0^{\pi/2} sen\, x\, dx = 1$, pues $d(-\cos x)/dx = sen\, x$ y entonces $\int_0^{\pi/2} sen\, x\, dx = -\cos(\pi/2) - (-\cos(0)) = 1$. El lector debería estar familiarizado con estas ideas. ♦

Recordemos la idea intuitiva que hay detrás del teorema **4.8.8**; para ello definamos

$$F(x) = \int_a^x f(t)\,dt$$

y supongamos que $f \geq 0$, para mayor sencillez; de este modo, F es el área debajo de la gráfica de f, entre a y x. El hecho de que $F' = f$ proviene de que $f(x)$ es la tasa con la que crece dicha área. Esto debería ser claro, pues $F(x + \Delta x) - F(x) \approx f(x)\Delta x$ (figura 4.8-2).

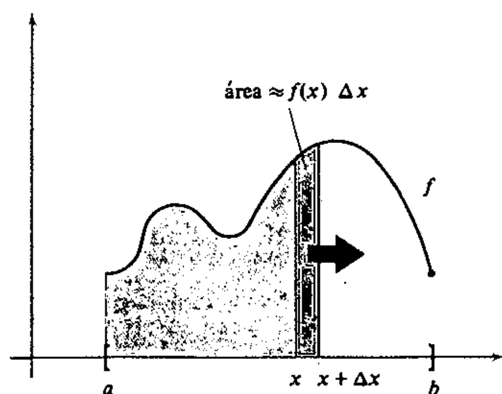


FIGURA 4.8-2 La geometría del teorema fundamental del cálculo

Suponemos que el lector conoce, o va a repasar, las técnicas básicas de integración que se obtienen de este teorema, como el método de sustitución (la regla de la cadena)

y la integración por partes (regla del producto). Daremos por sentado que así es en el resto del libro.

4.8.10 Ejemplo *Sea $F(x) = \int_a^x f(t)\,dt$. ¿Es F derivable si f es simplemente integrable (Riemann)?*

Solución No, la continuidad en el teorema fundamental no se puede reemplazar por la integrabilidad. Por ejemplo, sea $f(x) = 0$ en $[0, 1]$ y $f(x) = 1$ en $]1, 2]$. Entonces $F(x) = 0$ en $[0, 1]$ y $F(x) = x - 1$ en $]1, 2]$, que es continua pero no es derivable en $x = 1$ (véase la figura 4.8-3). ♦

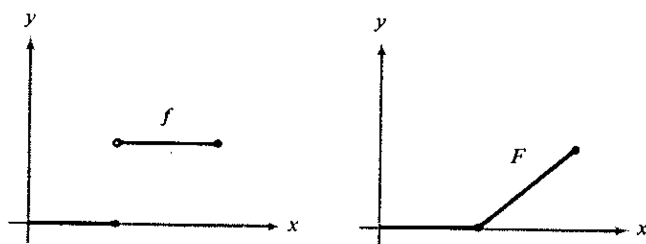


FIGURA 4.8-3 La integral "indefinida" F de f no tiene por qué ser derivable

Ejercicios de §4.8

1. Muéstrese directamente de la definición que $\int_a^b dx = b - a$.
2. Evalúese $\int_0^3 (x + 5)dx$ por medio de la definición de la integral.
3. Muéstrese directamente de la definición que si $f: [a, b] \to \mathbb{R}$ y P es cualquier partición de $[a, b]$, $U(f, P) \geq L(f, P)$.
4. Sea $f: [a, b] \subset \mathbb{R} \to \mathbb{R}$ una función integrable tal que $f \leq M$. Demuéstrese que

$$\int_a^b f(x)\,dx \leq (b - a)M.$$

5. Evalúese

$$\int_0^5 xe^{3x^2}\,dx.$$

6. Evalúese

$$\int_0^1 (x+2)^9\,dx.$$

7. Sea $f : [0, 1] \to \mathbb{R}$ tal que $f(x) = 1$ si $x = 1/n$, con n entero y $f(x) = 0$ en los demás casos.

a. Demuéstrese que f es integrable.

b. Muéstrese que $\int_0^1 f(x)\,dx = 0$.

8. Sea $f: [a, b] \to \mathbb{R}$ integrable Riemann y $|f(x)| \le M$. Sea $F(x) = \int_a^x f(t)\,dt$. Demuéstrese que $|F(y) - F(x)| \le M|y - x|$ y dedúzcase de ello que F es continua. ¿Concuerda esto con el ejemplo **4.8.10**?

Demostraciones de los teoremas del capítulo 4

4.1.4 Teorema *Sea $f : A \subset M \to N$ una transformación. Entonces las siguientes afirmaciones son equivalentes:*

i. *f es continua en A.*

ii. *Para cada sucesión convergente $x_k \to x_0$ en A, tenemos $f(x_k) \to f(x_0)$.*

iii. *Para cada conjunto abierto U en N, $f^{-1}(U) \subset A$ es abierto relativo a A; es decir, $f^{-1}(U) = V \cap A$ para algún conjunto abierto V.*

iv. *Para cada conjunto cerrado $F \subset N$, $f^{-1}(F) \subset A$ es cerrado relativo a A; es decir, $f^{-1}(F) = G \cap A$ para algún conjunto cerrado G.*

Demostración El patrón de la demostración será **i** $\Rightarrow$ **ii** $\Rightarrow$ **iv** $\Rightarrow$ **iii** $\Rightarrow$ **i.**

i $\Rightarrow$ **ii** Supongamos que $x_k \to x_0$. Para mostrar que $f(x_k) \to f(x_0)$, sea $\varepsilon > 0$; debemos encontrar un entero N tal que $k \ge N$ implique $\rho(f(x_k), f(x_0)) < \varepsilon$. Para ello, elegimos $\delta > 0$ de modo que $d(x, x_0) < \delta$ implique $\rho(f(x), f(x_0)) < \varepsilon$. La existencia de tal δ está garantizada por la continuidad de f. Entonces, elegimos N tal que $k \ge N$ implique $d(x_k, x_0) < \delta$. Esta elección de N produce la conclusión deseada.

ii⇒iv Sea $F \subset N$ un conjunto cerrado. Para mostrar que $f^{-1}(F)$ es cerrado en A, observemos que un conjunto B es cerrado relativo a A sii para cada sucesión $x_k \in B$ que converja a un punto $x \in A$, necesariamente ocurre que $x \in B$. Sea $x_k \in f^{-1}(F)$ tal que $x_k \to x$, donde $x \in A$. Debemos mostrar que $x \in f^{-1}(F)$. Por **ii**, $f(x_k) \to f(x)$, y como $f(x_k) \in F$ y F es cerrado, concluimos que $f(x) \in F$. Así, $x \in f^{-1}(F)$.

iv⇒iii Si U es abierto, sea $F = N\backslash U$, que es cerrado. Entonces, por **iv**, $f^{-1}(F) = G \cap A$ para algún conjunto cerrado G. Así, $f^{-1}(U) = A \cap (M \setminus G)$ y entonces $f^{-1}(U)$ es abierto relativo a A.

iii⇒i Dados $x_0 \in A$ y $\varepsilon > 0$, debemos encontrar δ tal que $d(x, x_0) < \delta$ implique $\rho(f(x), f(x_0)) < \varepsilon$. Puesto que $D(f(x_0), \varepsilon)$ es un conjunto abierto, $f^{-1}(D(f(x_0), \varepsilon))$ es abierto, por **iii**. Así, por la definición de conjunto abierto y el hecho de que $x_0 \in f^{-1}(D(f(x_0), \varepsilon))$, existe $\delta > 0$ tal que $D(x_0, \delta) \cap A \subset f^{-1}(D(f(x_0), \varepsilon))$. Ésta es otra forma de decir que

$$(x \in A \quad \text{y} \quad d(x, x_0) < \delta) \Rightarrow (\rho(f(x), f(x_0)) < \varepsilon). \quad \blacksquare$$

Para ganar soltura con estos conceptos, el lector debería intentar demostrar (directamente) otras implicaciones de este teorema; por ejemplo, **i ⇒ iii**, o **ii ⇒ i**.

4.2.1 Teorema *Supongamos que $f: M \to N$ es continua y que $K \subset M$ es conexo. Entonces $f(K)$ es conexo. Análogamente, si K es conexo por arcos, también lo es $f(K)$.*

Demostración *Supongamos* que $f(K)$ no es conexo. Por definición, podemos escribir $f(K) \subset U \cup V$, donde $U \cap V \cap f(K) = \varnothing$, $U \cap f(K) \neq \varnothing$, $V \cap f(K) \neq \varnothing$ y U, V son conjuntos abiertos. Ahora bien, $f^{-1}(U) = U' \cap A$ para algún conjunto abierto U' y, análogamente, $f^{-1}(V) = V' \cap A$ para algún conjunto abierto V'. Las condiciones sobre U, V implican que $U' \cap V' \cap K = \varnothing$, $K \subset U' \cup V'$, $U' \cap K \neq \varnothing$, y $V' \cap K \neq \varnothing$. Luego, K no es conexo, lo que demuestra nuestra primera afirmación.

Para la segunda parte, sean $z = f(x)$, $w = f(y)$ dos puntos en $f(K)$, donde $x, y \in K$. Sea $c(t)$ una curva continua que une x e y. Entonces $d(t) = f(c(t))$ es una curva continua que une z con w. Es continua, pues si $t_k \to t_0$, entonces $c(t_k) \to c(t_0)$, ya que c es continua; en consecuencia, $f(c(t_k)) = d(t_k) \to f(c(t_0)) = d(t_0)$, debido a que f es continua. Así, d es continua y, por lo tanto, $f(K)$ es conexo por arcos. $\blacksquare$

4.2.2 Teorema *Supongamos que $f: M \to N$ es continua y que $B \subset M$ es compacto. Entonces $f(B)$ es compacto.*

Demostración Sea y_k una sucesión en $f(B)$. Por el teorema **3.1.3**, debemos mostrar que y_k tiene una subsucesión convergente a un punto en $f(B)$. Sea $y_k = f(x_k)$ para $x_k \in B$. Como B es compacto, existe una subsucesión convergente, digamos, $x_{k_n} \to x$, con $x \in B$. Entonces, por el teorema **4.1.4ii**, $f(x_{k_n}) \to f(x)$, luego $f(x_{k_n})$ es una subsucesión convergente de y_k. ■

4.3.1 Teorema *Sean M, N y P espacios métricos y supongamos que* $f : A \subset M \to N$ *y* $g : B \subset N \to P$ *son transformaciones continuas tales que* $f(A) \subset B$. *Entonces* $g \circ f : A \subset M \to P$ *es continua.*

Demostración Sea $U \subset P$ un conjunto abierto. Entonces $(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U))$. Ahora, $g^{-1}(U) = U' \cap B$ para algún conjunto abierto U' y $f^{-1}(U' \cap B) = f^{-1}(U')$, pues $f(A) \subset B$. Como f es continua, $f^{-1}(U') = U'' \cap A$ para U'' abierto. Así, $g \circ f$ es continua, por el teorema **4.1.4.** ■

Las demás condiciones del teorema **4.1.4** se podrían usar también para demostrar el teorema **4.3.1.** En vez de demostrar el teorema **4.3.2**, nos restringiremos a demostrar su corolario. El caso general es similar; la única complejidad está en la notación.

4.3.3 Corolario *Sean* $A \subset M$ *y* $x_0 \in A$ *un punto de acumulación de A.*

i. *Sean* $f : A \to V$ *y* $g : A \to V$ *continuas en* x_0; *entonces la suma* $f + g : A \to V$ *es continua en* x_0.

ii. *Sean* $f : A \to \mathbb{R}$ *y* $g : A \to V$ *continuas en* x_0; *entonces el producto* $f \cdot g : A \to V$ *es continuo en* x_0.

iii. *Sean* $f : A \to \mathbb{R}$ *y* $g : A \to V$ *continuas en* x_0, *con* $f(x_0) \neq 0$; *entonces f es distinta de cero en una vecindad U de* x_0 *y el cociente* $g/f : U \to V$ *es continuo en* x_0.

Demostración

i. Sea $x_0 \in A$ y supongamos que $\varepsilon > 0$ está dado. Elegimos $\delta_1 > 0$ tal que $d(x, x_0) < \delta_1$ implica $d(f(x), f(x_0)) < \varepsilon/2$ y $\delta_2 > 0$ tal que $d(x, x_0) < \delta_2$ implica $d(g(x), g(x_0)) < \varepsilon/2$. Sea δ el mínimo de δ_1, δ_2. Por lo tanto, si $d(x, x_0) < \delta$, la desigualdad triangular implica que

$$\begin{aligned} \|(f+g)(x) - (f+g)(x_0)\| &= \|f(x) - f(x_0) + g(x) - g(x_0)\| \\ &\leq \|f(x) - f(x_0)\| + \|g(x) - g(x_0)\| \\ &\leq \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

ii. Sea $x_0 \in A$ y supongamos que $\varepsilon > 0$. Elijamos δ_1 tal que $d(x, x_0) < \delta_1$ implique $|f(x) - f(x_0)| < \varepsilon/2(\|g(x_0)\|) + 1)$ y $|f(x)| \leq |f(x_0)| + 1$ (¿por qué es posible esto?). Además, cojamos δ_2 tal que $d(x, x_0) < \delta_2$ implique $\|g(x) - g(x_0)\| < \varepsilon/[2(|f(x_0)| + 1)]$. Entonces, si $\delta = \text{mín}(\delta_1, \delta_2)$, $d(x, x_0) < \delta$ implica (por la desigualdad triangular)

$$\begin{aligned}\|fg(x) - fg(x_0)\| &= \|f(x)g(x) - f(x)g(x_0) + f(x)g(x_0) - f(x_0)g(x_0)\| \\ &\leq |f(x)|\,\|g(x) - g(x_0)\| + |f(x) - f(x_0)|\,\|g(x_0)\|\end{aligned}$$

(puesto que $\|av\| = |a|\,\|v\|$ para $v \in V$ y $a \in \mathbb{R}$). Continuando con esta línea de razonamiento obtenemos

$$\|fg(x) - fg(x_0)\| < \frac{(|f(x_0)| + 1)\varepsilon}{2(|f(x_0)| + 1)} + \frac{\|g(x_0)\|\varepsilon}{2(\|g(x_0)\| + 1)} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

iii. Por la demostración de **ii**, basta considerar el caso $1/f$, pues $g/f = g \cdot (1/f)$. Para mostrar que $1/f$ es continua, dado $x_0 \in A$, elegimos δ_1 tal que $|f(x) - f(x_0)| \leq (|f(x_0)|/2)$ para $\|x - x_0\| < \delta_1$. Esto es posible debido a la continuidad de f. De aquí se sigue que $|f(x)| \geq (|f(x_0)|/2)$. Ahora, dado $\varepsilon > 0$, elijamos δ_2 tal que $\|x - x_0\| < \delta_2$ implique

$$|f(x) - f(x_0)| < \frac{\varepsilon |f(x_0)|^2}{2}.$$

Tomando $\delta = \text{mín}(\delta_1, \delta_2)$, $\|x - x_0\| < \delta$ implica

$$\left|\frac{1}{f(x)} - \frac{1}{f(x_0)}\right| = \left|\frac{f(x_0) - f(x)}{f(x_0)\,f(x)}\right| \leq \frac{|f(x) - f(x_0)|}{|f(x_0)|^2/2} < \varepsilon.$$

Esto muestra que $1/f(x)$ es continua en x_0. ■

4.4.1 Teorema del máximo-mínimo *Sean (M, d) un espacio métrico, $A \subset M$ y $f: A \to \mathbb{R}$ una función continua. Sea $K \subset A$ un conjunto compacto. Entonces f está acotada en K; es decir, $B = \{f(x) \mid x \in K\} \subset \mathbb{R}$ es un conjunto acotado. Además, existen puntos $x_0, x_1 \in K$ tales que $f(x_0) = \inf(B)$ y $f(x_1) = \sup(B)$. Decimos que $\sup(B)$ es el máximo (absoluto) de f en K e $\inf(B)$ el mínimo (absoluto) de f en K.*

Demostración En primer lugar, B está acotado superiormente, ya que $B = f(K)$ es compacto por el teorema **4.2.2**, y, por tanto, es cerrado y acotado, por la definición de

compacidad. En segundo lugar, queremos obtener x_1 tal que $x_1 \in K$ y $f(x_1) = \sup(B)$. Ahora bien, como B es cerrado, $\sup(B) \in B = f(K)$ (véase el ejercicio 8, capítulo 2). Entonces, $\sup(B) = f(x_1)$ para algún $x_1 \in K$. El caso de $\inf(B)$ es similar. ■

Nota. También podemos demostrar el caso de $\inf(B)$ aplicando la demostración anterior (para el caso del supremo) a $-f$ y observando que el máximo de $-f$ es el mínimo de f cambiado de signo.

4.5.1 Teorema de los valores intermedios *Sean M un espacio métrico, $A \subset M$, y $f: A \to \mathbb{R}$ continua. Supongamos que $K \subset A$ es conexo y que $x,y \in K$. Para cada número $c \in \mathbb{R}$ tal que $f(x) < c < f(y)$, existe un punto $z \in K$ tal que $f(z) = c$.*

Demostración Supongamos que no existe tal z. Sean $U =]-\infty, c[= \{t \in \mathbb{R} \mid t < c\}$ y $V =]c, \infty[$. Es evidente que U y V son conjuntos abiertos. Como f es continua, tendremos que $f^{-1}(U) = U_0 \cap K$ para algún conjunto abierto U_0 y, análogamente, $f^{-1}(V) = V_0 \cap K$. Por la definición de U y V, tenemos que $U_0 \cap V_0 \cap K = \varnothing$ y por la hipótesis de que $\{z \in K \mid f(z) = c\} = \varnothing$, tenemos que $U_0 \cup V_0 \supset K$. Además, $U_0 \cap K \neq \varnothing$, pues $x \in U$, y $V_0 \cap K \neq \varnothing$ pues $y \in V_0$. Por lo tanto, K no es conexo, una contradicción. ■

4.6.2 Teorema de la continuidad uniforme *Sean $f: A \to N$ continua y $K \subset A$ un conjunto compacto. Entonces f es uniformemente continua en K.*

Demostración Dados $\varepsilon > 0$ y $x \in K$, sea δ_x tal que $d(x, y) < \delta_x$ implique $\rho(f(x), f(y)) < \varepsilon/2$. Los conjuntos $D(x, \delta_x/2)$ recubren K y son abiertos. Por lo tanto, existe un subrecubrimiento finito, digamos, $D(x_1, \delta_{x_1}/2), \ldots, D(x_N, \delta_{x_N}/2)$. Sea δ el mínimo de $\delta_{x_1}/2, \ldots, \delta_{x_N}/2$. Si $d(x, y) < \delta$, entonces existe x_i tal que $d(x, x_i) < \delta_{x_i}/2$ (puesto que los discos recubren K) y, por lo tanto, $d(x_i, y) \le d(x, x_i) + d(x, y) < \delta_{x_i}$. Así, por la elección de δ,

$$\rho(f(x), f(y)) \le \rho(f(x), f(x_i)) + \rho(f(x_i), f(y)) < \varepsilon/2 + \varepsilon/2 = \varepsilon. \quad ■$$

4.7.5 Teorema *Supongamos que f y g son derivables en x_0 y que $k \in \mathbb{R}$. Entonces kg, $f + g$, fg y f/g (si $g(x_0) \neq 0$) son derivables en x_0 y*

i. $(kf)'(x_0) = k(f'(x_0))$ ***regla de la multiplicación por una constante***

ii. $(f + g)'(x_0) = f'(x_0) + g'(x_0)$ ***regla para la suma***

iii. $(fg)'(x_0) = f(x_0)g'(x_0) + f'(x_0)g(x_0)$ ***regla del producto***

iv. $\left(\dfrac{f}{g}\right)'(x_0) = \dfrac{g(x_0)\, f'(x_0) - f(x_0)g'(x_0)}{[g(x_0)]^2}$ ***regla del cociente***

Demostración El resultado en cada caso se sigue de las propiedades de los límites. Por ejemplo, para demostrar **i,**

$$\begin{aligned}(kf)'(x_0) &= \lim_{x\to x_0} \frac{kf(x) - kf(x_0)}{x - x_0} = \lim_{x\to x_0} k\left[\frac{f(x) - f(x_0)}{x - x_0}\right] \\ &= k \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} = kf'(x_0).\end{aligned}$$

Los lectores deberían probar por sí mismos los demás apartados, consultando un texto de cálculo en caso necesario. ■

4.7.6 Regla de la cadena *Si f es derivable en x_0 y g es derivable en $f(x_0)$, entonces $g \circ f$ es derivable en x_0 y*

$$(g \circ f)'(x_0) = g'(f(x_0))\, f'(x_0).$$

Demostración Sean $y = f(x)$ y $z = g(y)$. El argumento usual, a saber,

$$\begin{aligned}\frac{dz}{dx} &= \lim_{\Delta x\to 0} \frac{\Delta z}{\Delta x} = \lim_{\Delta x\to 0} \left(\frac{\Delta z}{\Delta y}\frac{\Delta y}{\Delta x}\right) \\ &= \left(\lim_{\Delta x\to 0} \frac{\Delta z}{\Delta y}\right)\left(\lim_{\Delta x\to 0} \frac{\Delta y}{\Delta x}\right) \\ &= \left(\lim_{\Delta y\to 0} \frac{\Delta z}{\Delta y}\right)\left(\lim_{\Delta x\to 0} \frac{\Delta y}{\Delta x}\right) = \frac{dz}{dy}\frac{dy}{dx},\end{aligned}$$

tiene el defecto de que Δy podría anularse aunque Δx no lo haga.

Aquí presentamos un argumento más cuidadoso que evita estos problemas con los denominadores. La definición de derivabilidad de una función $k(u)$ en u se puede escribir como sigue: dado $\varepsilon > 0$, existe $\delta > 0$ tal que $|\Delta u| < \delta$ implica

$$|k(u + \Delta u) - k(u) - k'(u)\Delta u| < \varepsilon|\Delta u|.$$

Dado $\varepsilon > 0$, debemos encontrar $\delta > 0$ tal que

$$|g(f(x + \Delta x)) - g(f(x)) - g'(f(x))\, f'(x)\Delta x| < \varepsilon\, \Delta x \qquad (1)$$

si $|\Delta x| < \delta$. Aquí hemos escrito x en vez de x_0 para simplificar la notación. Si sumamos y restamos $g'(f(x))\Delta f$, el miembro izquierdo de (1) se lee

$$|\Delta g - g'(f(x))\Delta f + g'(f(x))\Delta f - g'(f(x))\, f'(x)\Delta x|$$
$$\leq |\Delta g - g'(f(x))\Delta f| + |g'(f(x))|\, |\Delta f - f'(x)\Delta x| \tag{2}$$

donde

$$\Delta g = g(f(x + \Delta x)) - g(f(x)) = g(f(x) + \Delta f) - g(f(x))$$

y

$$\Delta f = f(x + \Delta x) - f(x).$$

Como f tiene la propiedad de Lipschitz en x (véanse los comentarios posteriores a **4.7.2**), $|\Delta f| < M|\Delta x|$ si $|\Delta x| < \delta$ para algún $c > 0$. Como g es derivable en $f(x)$, dado $\varepsilon_1 > 0$, existe $\delta_1 > 0$ tal que si $|\Delta f| < \delta_1$, entonces $|\Delta g - g'(f(x))\Delta f| < \varepsilon_1 |\Delta f|$.

Así pues, hemos de elegir δ tal que $|\Delta f| < \delta_1$; es decir, $\delta \leq \delta_1/M$. Entonces, el miembro derecho de (2) está acotado superiormente por

$$\varepsilon_1|\Delta f| + |g'(f(x))|\Delta f - f'(x)\Delta x| \leq \varepsilon_1 M|\Delta x| + N|\Delta f - f'(x)\Delta x| \tag{3}$$

donde $N = |g'(f(x))|$. Dado $\varepsilon_2 > 0$, elijamos δ_2 tal que $|\Delta f - f'(x)\Delta x| < \varepsilon_2|\Delta x|$ si $|\Delta x| < \delta_2$. Entonces, el miembro izquierdo de (3) está acotado superiormente por

$$(\varepsilon_1 M + \varepsilon_2 N)|\Delta x|. \tag{4}$$

Ahora está claro cómo debemos elegir las cosas: dado $\varepsilon > 0$, sean $\varepsilon_1 = \varepsilon/2M$, $\varepsilon_2 = \varepsilon/2N$ y δ_1, δ_2 las deltas correspondientes. Sea $\delta = \text{mín}(\delta_1/M, \delta_2)$. Entonces (4) muestra la validez de (1). ■

4.7.8 Teorema *Supongamos que f es derivable en x_0.*

i. *Si f es creciente en x_0, entonces $f'(x_0) \geq 0$.*

ii. *Si f es decreciente en x_0, entonces $f'(x_0) \leq 0$.*

iii. *Si $f'(x_0) > 0$, entonces f es estrictamente creciente en x_0.*

iv. *Si $f'(x_0) < 0$, entonces f es estrictamente decreciente en x_0.*

Demostración **i** y **ii** se siguen de **iv** y **iii** (¿por qué?), y así lo ilustramos dando la demostración de **iii**. Así pues, supongamos que $f'(x_0) > 0$. Por definición,

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0).$$

Sea $\varepsilon = f'(x_0)$. Existe una vecindad N de x_0 tal que para todo $x \in N$ tal que $x \neq x_0$ se tiene

$$\left|\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0)\right| < f'(x_0),$$

o, de forma equivalente,

$$0 < \frac{f(x) - f(x_0)}{x - x_0} < 2f'(x_0).$$

Así,

$$\frac{f(x) - f(x_0)}{x - x_0}$$

es positivo para $x \in N$, $x \neq x_0$, de modo que si $x > x_0$, entonces $x - x_0 > 0$ y, por lo tanto, $f(x) - f(x_0) > 0$ (es decir, $f(x) > f(x_0)$); y si $x < x_0$, entonces $x - x_0 < 0$ y, por lo tanto, $f(x) - f(x_0) < 0$ (es decir, $f(x) < f(x_0)$). ■

4.7.14 Proposición *Supongamos que f es continua en $[a, b]$ y derivable en $]a, b[$.*

i. *Si $f'(x) \geq 0$ para todo $x \in\,]a, b[$, entonces f es creciente en $[a, b]$.*

ii. *Si $f'(x) \leq 0$ para todo $x \in\,]a, b[$, entonces f es decreciente en $[a, b]$.*

iii. *Si $f'(x) > 0$ para todo $x \in\,]a, b[$, entonces f es estrictamente creciente en $[a, b]$.*

iv. *Si $f'(x) < 0$ para todo $x \in\,]a, b[$, entonces f es estrictamente decreciente en $[a, b]$.*

v. *Si $f'(x) = 0$ para todo $x \in\,]a, b[$, entonces $f(x) = k$ (k es una constante) para todo $x \in [a, b]$.*

Demostración En cada caso, el resultado se sigue del teorema del valor medio. Ya hemos demostrado **i** en el texto; la demostración de los demás casos es similar. ■

4.7.15 Teorema de la función inversa *Supongamos que $f:]a, b[\to \mathbb{R}$ y que, o bien $f'(x) > 0$ para todo $x \in\,]a, b[$, o bien $f'(x) < 0$ para todo $x \in\,]a, b[$. Entonces $f:]a, b[\to \mathbb{R}$ es una biyección sobre su recorrido, f^{-1} es derivable en su dominio y $(f^{-1})'(y) = 1/f'(x)$, donde $f(x) = y$.*

Demostración Como f es monótona por la proposición anterior, f^{-1} existe; por **4.2.1** y **4.5.3**, su dominio es un intervalo. Primero probaremos que f^{-1} es continua. Supongamos que $f'(x) > 0$, de modo que f es estrictamente creciente (se puede hacer una demostración similar si f es estrictamente decreciente). Si $U \subset\,]a, b[$ es un conjunto abierto,

sea $y \in f(U)$, de modo que $y = f(x)$ para algún $x \in U$ y existe un intervalo abier-to $]x_1, x_2[\subset U$ tal que $x \in]x_1, x_2[$. Ahora, $f(x_1) < f(x_2)$ y $f(]x_1, x_2[) \subset]f(x_1), f(x_2)[$, pues f es estrictamente creciente. Para cualquier c tal que $f(x_1) < c < f(x_2)$, existe $z \in]x_1, x_2[$ que satisface $f(z) = c$, por el teorema de los valores intermedios. Así, $f(]x_1, x_2[) =]f(x_1), f(x_2)[\subset f(U)$ e $y = f(x) \in]f(x_1), f(x_2)[$. Hemos probado que $(f^{-1})^{-1}(U) = f(U)$ es abierto, por lo tanto, el teorema **4.1.4** implica que f^{-1} es continua.

Ahora escribamos $y = f(x)$ e $y_0 = f(x_0)$, de modo que $x = f^{-1}(y)$ y $x_0 = f^{-1}(y_0)$. Como f^{-1} es continua, $\lim_{y\to y_0} x = x_0$. Entonces,

$$(f^{-1})'(y_0) = \lim_{y\to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \lim_{y\to y_0} \frac{x - x_0}{f(x) - f(x_0)}$$
$$= \lim_{y\to y_0} \frac{1}{\dfrac{f(x) - f(x_0)}{x - x_0}} = \frac{1}{f'(x_0)}. \quad \blacksquare$$

4.7.16 Proposición *Supongamos que f es continua en $[a, b]$, que es derivable dos veces en $]a, b[$, y que $x_0 \in]a, b[$.*

i. *Si $f'(x_0) = 0$ y $f''(x_0) > 0$, entonces x_0 es un mínimo local estricto de f.*

ii. *Si $f'(x_0) = 0$ y $f''(x_0) < 0$, entonces x_0 es un máximo local estricto de f.*

Demostración

i. Si $f''(x_0) > 0$, entonces $f'(x)$ es creciente en x_0, por lo que existe $\delta > 0$ tal que $f'(x) < 0$ si $x_0 - \delta < x < x_0$ y $f'(x) > 0$ si $x_0 < x < x_0 + \delta$. Así, por **4.7.14**, $f(x) > f(x_0)$ si $x_0 - \delta < x < x_0$ y $f(x) > f(x_0)$ si $x_0 < x < x_0 + \delta$.

ii. La demostración es similar. $\blacksquare$

4.8.4 Teorema

i. *Si $f: [a, b] \to \mathbb{R}$ está acotada y es continua en todos los puntos de $[a, b]$ excepto en un número finito de ellos, entonces es integrable en $[a, b]$.*

ii. *Cualquier función creciente o decreciente en $[a, b]$ es integrable en $[a, b]$.*

Primero demostraremos lo siguiente:

Desigualdad entre las integrales superior e inferior *Para una función acotada $f: [a, b] \to \mathbb{R}$, $\underline{\int_a^b} f \le \overline{\int_a^b} f$.*

Demostración Si P y P' son particiones de $[a, b]$ tales que $P \subset P'$, entonces P' es un ***refinamiento*** de P. Para obtener nuestra desigualdad, primero demostraremos el siguiente resultado:

Lema *Si P' es un refinamiento de P, entonces $L(f, [a, b], P') \geq L(f, [a, b], P)$ y $U(f, [a, b], P') \leq U(f, [a, b], P)$.*

Demostración Sean $P_0 = P$ y $P_1 = P \cup \{a_1\}$, donde $a_1 \in P' \setminus P$, y sea $P_i = P_{i-1} \cup \{a_i\}$, donde $a_i \in P'/P_{i-1}$, de modo que, para algún t, $P = P_0 \subset P_1 \subset P_2 \subset \cdots P_t = P'$ y P_{i+1} tiene un punto más que P_i. Si $P_i = \{x_0, x_1, \ldots, x_n\}$ con $x_j < x_{j+1}$ y $x_{k-1} < a_{i+1} < x_k$, entonces

$$\begin{aligned} L(f, [a,b], P_i) = {} & m_1(x_1 - x_0) + m_2(x_2 - x_1) + \cdots + m_{k-1}(x_{k-1} - x_{k-2}) \\ & + m_k(x_k - x_{k-1}) \\ & + m_{k+1}(x_{k+1} - x_k) + \cdots + m_n(x_n - x_{n-1}) \end{aligned}$$

y

$$\begin{aligned} L(f, [a,b], P_{i+1}) = {} & m_1(x_1 - x_0) + m_2(x_2 - x_1) + \cdots + m_{k-1}(x_{k-1} - x_{k-2}) \\ & + l_1(a_{i+1} - x_{k-1}) + l_2(x_k - a_{i+1}) \\ & + m_{k+1}(x_{k+1} - x_k) + \cdots + m_n(x_n - x_{n-1}) \end{aligned}$$

donde $m_j = \inf \{f(x) \mid x \in [x_{j-1}, x_j]\}$, $l_1 = \inf \{f(x) \mid x \in [x_{k-1}, a_{i+1}]\}$ y $l_2 = \inf \{f(x) \mid x \in [a_{k+1}, x_k]\}$. Las dos sumas son idénticas, excepto por que en la segunda suma, $m_k(x_k - x_{k-1})$ se reemplaza por $l_1(a_{i+1} - x_{k-1}) + l_2(x_k - a_{i+1})$. Como $m_k = \inf \{f(x) \mid x \in [x_{k-1}, x_k]\}$, m_k es una cota inferior de $\{f(x) \mid x \in [x_{k-1}, a_{i+1}]\}$ y $\{f(x) \mid x \in [a_{i+1}, x_k]\}$. Así pues, $m_k \leq l_1$ y $m_k \leq l_2$, y

$$\begin{aligned} l_1(a_{i+1} - x_{k-1}) + l_2(x_k - a_{i+1}) &\geq m_k(a_{i+1} - x_{k-1}) + m_k(x_k - a_{i+1}) \\ &= m_k(x_k - x_{k-1}). \end{aligned}$$

luego, $L(f, [a, b], P_{i+1}) \geq L(f, [a, b], P_i)$. Como esta desigualdad es válida para $i = 0, 1, \ldots, t-1$, tenemos que $L(f, [a, b], P') \geq L(f, [a, b], P)$. La demostración de este resultado para el caso de las sumas superiores es similar. ▼

Para una partición P dada de $[a, b]$, tenemos que $M_i \geq m_i$, de modo que $L(f, [a, b], P) \leq U(f, [a, b], P)$. Ahora, sean P_1, P_2 dos particiones de $[a, b]$. Entonces, por el lema,

$$L(f, [a,b], P_1) \leq L(f, [a,b], P_1 \cup P_2) \leq U(f, [a,b], P_1 \cup P_2) \leq U(f, [a,b], P_2).$$

Así pues, una suma superior de cualquier partición es una cota superior para

$$\{L(f, [a, b], P) \mid P \text{ es una partición de } [a, b]\}.$$

Por lo tanto, para cualquier partición P, $\underline{\int_a^b} f \leq U(f, [a, b], P)$. Así, $\underline{\int_a^b} f$ es una cota inferior de $\{U(f, [a, b], P \mid P$ es una partición de $[a, b]\}$ y, por lo tanto,

$$\underline{\int_a^b} f \leq \overline{\int_a^b} f. \quad \blacksquare$$

Demostración de 4.8.4

$\underline{\int_a^b} f \geq L(f, [a, b], P)$ para cualquier partición P de $[a, b]$ y

$$\overline{\int_a^b} f \leq U(f, [a,b], P).$$

Por lo tanto, para probar que $\underline{\int_a^b} f = \overline{\int_a^b} f$, basta demostrar que para cada $\varepsilon > 0$ existe una partición P de $[a, b]$ tal que $U(f, [a, b], P) - L(f, [a, b], P) < \varepsilon$.

i. Supongamos que f es continua en $[a, b]$ excepto en $a_1, a_2, \ldots, a_m$. Como f está acotada, existen números k_1, k_2 tales que $k_1 < f(x) < k_2$ para todo $x \in [a, b]$. Cojamos una partición P_1 tal que cada subintervalo que incluya algún a_i tenga longitud menor o igual que $\varepsilon/2m(k_2 - k_1)$. Como f es continua en la unión U de los demás subintervalos de P_1, f es uniformemente continua en U. Por lo tanto, existe δ tal que $x_1, x_2 \in U$ y $|x_1 - x_2| < \delta$ implican $|f(x_1) - f(x_2)| < \varepsilon/2(b - a)$. Refinamos P_1, en caso necesario, de modo que cada subintervalo de la partición refinada P que no contenga ningún a_i tenga longitud menor que δ. Ahora bien,

$$U(f, [a,b], P) - L(f, [a,b], P) = \sum_{j=1}^{n} (M_j - m_j)(x_j - x_{j-1}),$$

donde P está formada por $a = x_0 < x_1 < \cdots x_n = b$, $M_j = \sup \{f(x) \mid x \in [x_{j-1}, x_j]\}$ y $m_j = \inf \{f(x) \mid x \in [x_{j-1}, x_j]\}$. Para cada j, $M_j \leq k_2$ y $m_j \geq k_1$, pues k_2 y k_1 son las cotas superior e inferior de $f(x)$ en $[a, b]$. Estimemos primero la parte de la suma con los subintervalos que contienen las discontinuidades y después el resto de la suma.

$$\begin{aligned}
\sum_{\substack{j \text{ con algún} \\ a_i \in [x_{j-1}, x_j]}} (M_j - m_j)(x_j - x_{j-1}) &\leq \sum_{\substack{j \text{ con algún} \\ a_i \in [x_{j-1}, x_j]}} (k_2 - k_1)(x_j - x_{j-1}) \\
&= (k_2 - k_1) \sum_{\substack{j \text{ con algún} \\ a_i \in [x_{j-1}, x_j]}} \frac{\varepsilon}{2m(k_2 - k_1)} \\
&\leq (k_2 - k_1) \frac{\varepsilon}{2(k_2 - k_1)} = \frac{\varepsilon}{2},
\end{aligned}$$

puesto que a lo sumo existen m subintervalos que incluyen algún punto de discontinuidad. Para los $[x_{j-1}, x_j]$ donde f es continua, f alcanza un valor máximo M_j y un valor mínimo m_j; como el intervalo tiene longitud menor que δ y está contenido en U, $M_j - m_j < \varepsilon/2(b-a)$. Luego

$$\sum_{\substack{j \text{ con } f \text{ continuo} \\ \text{en } [x_{j-1},\, x_j]}} (M_j - m_j)(x_j - x_{j-1}) < \frac{\varepsilon}{2(b-a)} \sum_{\substack{j \text{ con } f \text{ continuo} \\ \text{en } [x_{j-1},\, x_j]}} (x_j - x_{j-1})$$

$$\leq \frac{\varepsilon}{2(b-a)} \cdot (b-a) = \frac{\varepsilon}{2},$$

puesto que la longitud total de los subintervalos no puede exceder la longitud de $[a, b]$. Así pues,

$$\begin{aligned} U(f,[a,b],P) - L(f,[a,b],P) &= \sum_{j=1}^{n} (M_j - m_j)(x_j - x_{j-1}) \\ &= \sum_{\substack{j \text{ con algún} \\ a_i \in [x_{j-1},\, x_j]}} (M_j - m_j)(x_j - x_{j-1}) \\ &\quad + \sum_{\substack{j \text{ con } f \text{ continuo} \\ \text{en } [x_{j-1},\, x_j]}} (M_j - m_j)(x_j - x_{j-1}) \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

Por lo tanto, f es integrable en $[a, b]$.

ii. Si f es creciente en $[a, b]$, entonces $m_i = f(x_{i-1})$ y $M_i = f(x_i)$. Así,

$$L(f,[a,b],P) = \sum_{i=1}^{n} f(x_{i-1})(x_i - x_{i-1})$$

y

$$U(f,[a,b],P) = \sum_{i=1}^{n} f(x_i)(x_i - x_{i-1}).$$

Elijamos una partición P tal que todos los subintervalos tengan longitud $(b-a)/n$. Entonces

$$\begin{aligned} U(f,[a,b],P) - L(f,[a,b],P) &= \frac{b-a}{n}\left(\sum_{i=1}^{n} f(x_i) - \sum_{i=1}^{n} f(x_{i-1})\right) \\ &= \frac{b-a}{n}(f(x_n) - f(x_0)) \\ &= \frac{b-a}{n}(f(b) - f(a)). \end{aligned}$$

Tomemos $n > (b-a)(f(b)-f(a))/\varepsilon$. Entonces

$$U(f,[a,b],P) - L(f,[a,b],P) < \frac{(b-a)(f(b)-f(a))}{(b-a)(f(b)-f(a))/\varepsilon} = \varepsilon$$

y, por lo tanto, f es integrable en $[a, b]$. La demostración para el caso de una función decreciente es similar. ■

4.8.5 Proposición

i. *Si f es integrable en $[a, b]$ y $k \in \mathbb{R}$, entonces kf es integrable en $[a, b]$ y* $\int_a^b kf = k\int_a^b f$.

ii. *Si f y g son integrables en $[a, b]$, entonces $f+g$ es integrable en $[a, b]$ y* $\int_a^b (f+g) = \int_a^b f + \int_a^b g$.

iii. *Si f y g son integrables en $[a, b]$ y $f(x) \leq g(x)$ para todo $x \in [a, b]$, entonces* $\int_a^b f \leq \int_a^b g$.

iv. *Si f es integrable en $[a, b]$ y en $[b, c]$, entonces f es integrable en $[a, c]$ y* $\int_a^c f = \int_a^b f + \int_b^c f$.

Demostración

i. Para una partición P de $[a, b]$, sean $m_i(kf) = \inf\{kf(x) \mid x \in [x_{i-1}, x_i]\}$ y $m_i(f) = \inf\{f(x) \mid x \in [x_{i-1}, x_i]\}$. Si $k \geq 0$, entonces $m_i(kf) = k \cdot m_i(f)$ y

$$\begin{aligned} L(kf,[a,b],P) &= \sum_{i=1}^{n} m_i(kf)(x_i - x_{i-1}) \\ &= \sum_{i=1}^{n} km_i(f)(x_i - x_{i-1}) \\ &= k\sum_{i=1}^{n} m_i(f)(x_i - x_{i-1}) = k \cdot L(f,[a,b],P). \end{aligned}$$

Puesto que esto es válido para cualquier partición P de $[a, b]$,

$$\underline{\int_a^b} kf = k\underline{\int_a^b} f = k\int_a^b f.$$

Análogamente

$$\overline{\int_a^b} kf = k\overline{\int_a^b} f = k\int_a^b f,$$

y, por lo tanto,

$$\int_a^b kf = \underline{\int_a^b} kf = \overline{\int_a^b} kf = k\int_a^b f.$$

Si $k < 0$, entonces $M_i(kf) = k \cdot m_i(f)$, de modo que

$$\begin{aligned} U(kf, [a,b], P) &= \sum_{i=1}^{n} m_i(kf)(x_i - x_{i-1}) = \sum_{i=1}^{n} km_i(f)(x_i - x_{i-1}) \\ &= k\sum_{i=1}^{n} m_i(f)(x_i - x_{i-1}) = k \cdot L(f, [a,b], P). \end{aligned}$$

Por lo tanto,

$$\overline{\int_a^b} kf = k\underline{\int_a^b} f = k\int_a^b f.$$

Análogamente,

$$\underline{\int_a^b} kf = k\overline{\int_a^b} f = k\int_a^b f$$

y entonces

$$\int_a^b kf = \underline{\int_a^b} kf = \overline{\int_a^b} kf = k\int_a^b f.$$

ii. Sea $m_i(f+g) = \inf\{(f+g)(x) \mid x \in [x_{i-1}, x_i]\}$. Como $f(x) + g(x) \geq m_i(f) + m_i(g)$ para todo $x \in [x_{i-1}, x_i]$, obtenemos $m_i(f+g) \geq m_i(f) + m_i(g)$, luego

$$\begin{aligned} L(f+g, [a,b], P) &= \sum_{i=1}^{n} m_i(f+g)(x_i - x_{i-1}) \\ &\geq \sum_{i=1}^{n} (m_i(f) + m_i(g))(x_i - x_{i-1}) \\ &= L(f, [a,b], P) + L(g, [a,b], P). \end{aligned}$$

Por lo tanto,

$$\underline{\int_a^b} (f+g) \geq \underline{\int_a^b} f + \underline{\int_a^b} g = \int_a^b f + \int_a^b g.$$

De forma similar,

$$\overline{\int_a^b} (f+g) \leq \overline{\int_a^b} f + \overline{\int_a^b} g = \int_a^b f + \int_a^b g.$$

Así que

$$\int_a^b f + \int_a^b g \leq \underline{\int_a^b} (f+g) \leq \overline{\int_a^b} (f+g) \leq \int_a^b f + \int_a^b g.$$

Por lo tanto,

$$\underline{\int_a^b} (f+g) = \overline{\int_a^b} (f+g) = \int_a^b f + \int_a^b g$$

y, en consecuencia,

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g.$$

iii. Como $f(x) \leq g(x)$ para todo $x \in [a, b]$, $m_i(f) \leq m_i(g)$ y $L(f, [a, b], P) \leq L(g, [a, b], P)$. Así pues,

$$\int_a^b f = \underline{\int_a^b} f \leq \underline{\int_a^b} g = \int_a^b g.$$

iv. Dadas las particiones $P_1 : a = x_0 < x_1 < \cdots < x_k = b$ de $[a, b]$ y $P_2 : b = x_k < x_{k+1} < \cdots < x_n = c$ de $[b, c]$, podemos formar una partición $P = P_1 \cup P_2$ de $[a, c]$. Ahora,

$$\begin{aligned} L(f,[a,c],P) &= \sum_{i=1}^{n} m_i(x_i - x_{i-1}) = \sum_{i=1}^{k} m_i(x_i - x_{i-1}) + \sum_{i=k+1}^{n} m_i(x_i - x_{i-1}) \\ &= L(f,[a,c],P_1) + L(f,[a,c],P_2). \end{aligned}$$

Así,

$$\begin{aligned} &\{L(f, [a, c], P) \mid P \text{ es una partición de } [a, c]\} \\ &\qquad \supset \{L(f,[a,c],P_1) + L(f,[a,c],P_2) \mid P_1 \\ &\qquad\qquad \text{es una partición de } [a, b] \text{ y } P_2 \text{ es una partición de } [b, c]\}. \end{aligned}$$

Por lo tanto,

$$\underline{\int_a^c} f \geq \underline{\int_a^b} f + \underline{\int_b^c} f = \int_a^b f + \int_b^c f.$$

De forma similar,

$$\overline{\int_a^c} f \leq \int_a^b f + \int_b^c f.$$

y entonces

$$\int_a^b f + \int_b^c f \le \underline{\int_a^c} f \le \overline{\int_a^c} f \le \int_a^b f + \int_b^c f$$

es decir,

$$\underline{\int_a^c} f = \overline{\int_a^c} f = \int_a^c f = \int_a^b f + \int_b^c f. \quad \blacksquare$$

4.8.8 Teorema fundamental del cálculo *Sea $f: [a, b] \to \mathbb{R}$ continua. Entonces f tiene una primitiva F y*

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Si G es otra primitiva de f, también tenemos que $\int_a^b f(x)\,dx = G(b) - G(a)$.

Demostración Definamos $F : [a, b] \to \mathbb{R}$ como $F(x) = \int_a^x f(y)\,dy$. Afirmamos que F es una primitiva de f. Sean $x \in]a, b[$ y $h > 0$ tales que $]x, x + h[\subset]a, b[$. Por **4.8.5iv**,

$$\frac{F(x + h) - F(x)}{h} = \frac{1}{h}\left(\int_a^{x+h} f(y)\,dy - \int_a^x f(y)\,dy\right) = \frac{1}{h}\left(\int_x^{x+h} f(y)\,dy\right).$$

Dado $\varepsilon > 0$, sea h tal que $|f(y) - f(x)| < \varepsilon$ para todo $y \in]x, x + h[$; tal h existe por la continuidad de f. Usando este hecho junto con

$$\int_x^{x+h} \frac{f(x)}{h}\,dy = f(x),$$

obtenemos

$$\left|\int_x^{x+h} \frac{f(y)}{h}\,dy - f(x)\right| = \left|\int_x^{x+h} \left(\frac{f(y) - f(x)}{h}\right)\,dy\right|$$

$$\le \int_x^{x+h} \left|\frac{f(y) - f(x)}{h}\right|\,dy < \frac{\varepsilon h}{h} = \varepsilon.$$

Por lo tanto,

$$\lim_{h \to 0^+} \frac{F(x + h) - F(x)}{h} = f(x) \quad \text{(el límite cuando } h > 0\text{)}.$$

De forma similar,

$$\lim_{h\to 0^-} \frac{F(x) - F(x-h)}{h} = f(x) \quad \text{(el límite cuando } h < 0\text{).}$$

Por lo tanto, F' existe y $F'(x) = f(x)$. Por la definición de F, obtenemos

$$\int_a^b f(y)\,dy = F(b) - F(a).$$

Se puede ver fácilmente que la función F es continua en a y b. Ahora, sea G cualquier primitiva de f; mostraremos que $G = F +$ constante. Como $G'(x) = F'(x) = f(x)$ para todo $x \in]a, b[$, tendremos que $(G - F)'(x) = 0$ para todo $x \in]a, b[$. Por **4.7.12**, si una función tiene derivada nula en un intervalo, es constante en ese intervalo. Por lo tanto, $G = F +$ constante y, en consecuencia,

$$\int_a^b f(y)\,dy = G(b) - G(a)$$

para cualquier primitiva G de f. ■

Ejemplos resueltos del capítulo 4

Ejemplo 4.1 *Sea $f : A \to \mathbb{R}^m$ dada por*

$$f(x) = (f_1(x), \ldots, f_m(x)).$$

Muéstrese que f es continua sii cada componente f_i es continua, $i = 1, \ldots, m$.

Solución Sea f continua. Si $x_k \to x$ en A, debemos probar que $f_i(x_k) \to f_i(x)$ para cada i. Pero esto es consecuencia del hecho de que $f(x_k) \to f(x)$ y una sucesión en $\mathbb{R}^m$ (en este caso, $f(x_k)$) converge sii sus sucesiones componentes convergen (véase §**2.7**). El mismo razonamiento demuestra el recíproco ♦

Ejemplo 4.2 *Sea $f : A \subset M \to N$ continua. Si $K \subset A$ es un conjunto conexo, muéstrese que la gráfica de f, $\{(x, f(x)) \mid x \in K\}$, es conexa en $M \times N$.*

Solución Considérese la transformación $g : K \subset M \to M \times N$ definida por $g(x) = (x, f(x))$. Por el ejemplo anterior, g es continua. Pero $g(K)\{(x, f(x)) \mid x \in K\}$, y la imagen de un conjunto conexo es conexa, por el teorema **4.2.1**. ♦

Ejemplo 4.3 *Sea $f: A \to \mathbb{R}^m$ continua en $x_0 \in A$, A abierto y $f(x_0) \neq 0$. Muéstrese que f es distinta de cero en alguna vecindad de x_0.*

Solución Dado $\varepsilon > 0$, existe una vecindad U de x_0 tal que $\|f(x) - f(x_0)\| < \varepsilon$ para todo $x \in U$, por la definición de continuidad. Para nuestros fines, elegimos $\varepsilon = \|f(x_0)\|$. Entonces $\|f(x) - f(x_0)\| < \|f(x_0)\|$ implica que $f(x) \neq 0$, pues $f(x) = 0$ implicaría la contradicción $\|-f(x_0)\| < \|f(x_0)\|$. Por lo tanto, en la vecindad U determinada por $\varepsilon = \|f(x_0)\|$, f es distinta de cero. ♦

Ejemplo 4.4 *Muéstrese que una aplicación lineal $L : \mathbb{R}^n \to \mathbb{R}^m$ es continua.*

Solución Mostraremos que para esta aplicación lineal $L : \mathbb{R}^n \to \mathbb{R}^m$ podemos determinar un número M tal que $\|L(x)\| \leq M\|x\|$ para todo $x \in \mathbb{R}^n$. Entonces $\|x - x_0\| < \varepsilon/M$ implica que $\|L(x) - L(x_0)\| = \|L(x - x_0)\| \leq M\|x - x_0\| < \varepsilon$, lo que demuestra que L es continua.

Sea $M_1 = \sup\{\|L(e_1)\|, \ldots, \|L(e_n)\|\}$ donde $e_1, \ldots, e_n$ es la base canónica de $\mathbb{R}^n$. Sea $x = (x_1, \ldots, x_n)$ y usemos la desigualdad triangular,

$$\begin{aligned} \|L(x)\| &= \|x_1 L(e_1) + \cdots + x_n L(e_n)\| \leq |x_1|\,\|L(e_1)\| + \cdots + |x_n|\,\|L(e_n)\| \\ &\leq M_1(|x_1| + \cdots + |x_n|) \leq M_1 n\|x\|. \end{aligned}$$

Entonces, podemos tomar $M = nM_1$, de donde obtenemos nuestro resultado. ♦

Ejemplo 4.5 Una ***aplicación multilineal*** L de $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \cdots \times \mathbb{R}^{n_k}$ a $\mathbb{R}^m$ *es una transformación tal que para cada r, $1 \leq r \leq k$, cumple*

$$L(a_1, a_2, \ldots, a_r + \lambda b_r, \ldots, a_k) = L(a_1, \ldots, a_r, \ldots, a_k) + \lambda L(a_1, a_2, \ldots, b_r, \ldots, a_k),$$

donde $a_i \in \mathbb{R}^{n_i}$, $b_r \in \mathbb{R}^{n_r}$ y $\lambda \in \mathbb{R}$. Muéstrese que una aplicación multilineal es continua.

Solución Sea $e_1, \ldots, e_n$ la base canónica de $\mathbb{R}^n$, y, para cada $x \in \mathbb{R}^n$, sea $x = (x^1, \ldots, x^n) = \sum_{i=1}^{n} x^i e_i$. Definimos los elementos de $\mathbb{R}^m$, $\alpha_{i_1, \ldots, i_k}$ para los enteros i_j tales que $1 \leq i_j \leq n_j$, $j = 1, \ldots, k$ como $\alpha_{i_1, \ldots, i_k} = L(e_{i_1}, \ldots, e_{i_k})$. Afirmamos que

$$L(x_1, \ldots, x_k) = \sum_{i_1=1}^{n_1} \cdots \sum_{i_k=1}^{n_k} \alpha_{i_1, \ldots, i_k} x_1^{i_1} \cdots x_k^{i_k},$$

que es el análogo a la expresión de una transformación lineal en términos de una matriz.

En efecto, por la multilinealidad,

$$L(x_1, \ldots, x_k) = L\left(\sum_{i_1=1}^{n_1} x_1^{i_1} e_{i_1}, x_2, \ldots, x_k\right) = \sum_{i_1=1}^{n_1} x_1^{i_1} L(e_{i_1}, x_2, \ldots, x_k).$$

El resultado se obtiene al repetir la operación k veces.

Esta fórmula muestra claramente que L es continua, puesto que las funciones $x_1^{i_1}$ a $x_k^{i_k}$ son productos de funciones continuas y son, por tanto, continuas, y L es una suma de estas funciones.

Otra solución de este problema, que procede como en el ejemplo resuelto **4.4**, consiste en mostrar que existe una constante $M > 0$ tal que $\|L(x_1, \cdots, x_k)\| \leq M\|x_1\| \cdots \|x_k\|$, de lo cual puede deducirse de forma directa la continuidad. ♦

Ejemplo 4.6 *Muéstrese que $\int_0^a x^2 dx = a^3/3$ por medio del teorema fundamental. Verifíquese esta respuesta directamente, mostrando que para cualquier $\varepsilon > 0$ dado existe una partición P de $[0, a]$ tal que $U(x^2, P) - L(x^2, P) < \varepsilon$ y que*

$$\inf\{U(x^2, P) \mid P \text{ es cualquier partición}\} = \sup\{L(x^2, P) \mid P \text{ es cualquier partición}\} = \frac{a^3}{3}.$$

Solución La función F definida por $F(x) = x^3/3$ es una primitiva de $f(x) = x^2$, pues $F'(x) = x^2$. Así, por el teorema fundamental,

$$\int_0^a x^2\, dx = F(a) - F(0) = \frac{a^3}{3}.$$

Para verificar nuestra respuesta por medio de las sumas superiores e inferiores, separamos $[0, a]$ en los n subintervalos $[0, a/n], [a/n, 2a/n], \ldots, [(n-1)a/n, a]$. Si llamamos P a esta partición, entonces

$$U(x^2, P) = \sum_{k=1}^{n} \left(\frac{ka}{n}\right)^2 \frac{a}{n} = \left(\frac{a^3}{n^3}\right)\left(\sum_{k=0}^{n} k^2\right) = \left(\frac{a^3}{n^3}\right)\left(\frac{1}{6}\right) n(n+1)(2n+1)$$

(véase el ejercicio 41) y

$$\begin{aligned} L(x^2, P) &= \sum_{k=1}^{n} \left(\frac{(k-1)a}{n}\right)^2 \frac{a}{n} \\ &= \left(\frac{a^3}{3}\right)\left(\sum_{k=0}^{n-1} k^2\right) = \left(\frac{a^3}{n^3}\right)\left(\frac{1}{6}\right)(n-1)(n)(2n-1) \end{aligned}$$

Por lo tanto,

$$U(x^2, P) = \left(\frac{a^3}{3}\right)\left(1 + \frac{1}{n}\right)\left(1 + \frac{1}{2n}\right)$$

y

$$L(x^2, P) = \left(\frac{a^3}{3}\right)\left(1 - \frac{1}{n}\right)\left(1 - \frac{1}{2n}\right)$$

Si elegimos n suficientemente grande, obtendremos $U(x^2, P) - L(x^2, P) < \varepsilon$, y así, $\inf\{U(x^2, P) \mid P$ es cualquier partición$\}$ $\sup\{L(x^2, P) \mid P$ es cualquier partición$\} = a^3/3$. ♦

Ejemplo 4.7

a. *Formúlese una definición y algunas propiedades básicas de la integración de funciones de valores complejos en un intervalo.*

b. *Evalúese* $\int_0^\pi e^{ix}\, dx$.

Solución

a. Cualquier función $f: [a, b] \to C$ se puede separar en dos funciones de dos valores reales, f_1 y f_2, tales que $f(x) = f_1(x) + if_2(x)$ (es decir, definimos $f_1(x)$ como la parte real de $f(x)$ y $f_2(x)$ como su parte imaginaria). Entonces, planteamos la definición natural

$$\int_a^b f(x)\, dx = \int_a^b f_1(x)\, dx + i \int_a^b f_2(x)\, dx$$

si f_1 y f_2 son integrables en $[a, b]$. En este caso, f es ***integrable en*** $[a, b]$. Es fácil ver que

i. $\int_a^b cf(x)\, dx = c \int_a^b f(x)\, dx$ para cualquier número complejo c y cualquier función integrable f.

ii. $\int_a^b [f(x) + g(x)]\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx$ para sendas funciones integrables f y g; y denotando con una barra superior la conjugación compleja,

iii. $\int_a^b \overline{f(x)}\, dx = \overline{\int_a^b f(x)\, dx}$.

Con un poco más de esfuerzo, también se puede demostrar que

iv. $\left|\int_a^b f(x)\, dx\right| \leq \int_a^b |f(x)|\, dx$.

Análogamente podemos definir $f'(x) = f_1'(x) + if_2'(x)$, y en ese caso se cumplen las reglas usuales de las derivadas. Se deja al lector que formule el teorema fundamental del cálculo en el ejercicio 46.

b. Como $e^{ix} = \cos x + i \operatorname{sen} x$, obtenemos $\int_0^\pi e^{ix}\,dx = \int_0^\pi (\cos x + i \operatorname{sen} x)\,dx = \operatorname{sen} x\,|_0^\pi + i(-\cos x)\,|_0^\pi = 0 + 2i = 2i$. ♦

Ejercicios del capítulo 4

1. **a.** Demuéstrese directamente que la función $1/x^2$ es continua en $]0, \infty[$.

b. Una función constante $f: A \to \mathbb{R}^m$ es una función tal que $f(x) = f(y)$ para todo $x, y \in A$. Muéstrese que f es continua.

c. ¿Es continua la función $f(y) = 1/(y^4 + y^2 + 1)$? Justifíquese la respuesta.

2. **a.** Demuéstrese que si $f: A \to \mathbb{R}^m$ es continua y $B \subset A$, entonces la restricción $f\,|\,B$ es continua.

b. Encuéntrese una función $g : A \to \mathbb{R}$ y un conjunto $B \subset A$ tales que $g\,|\,B$ sea continua pero que g no sea continua en ningún punto de A.

3. **a.** Si $f: \mathbb{R} \to \mathbb{R}$ es continua y $K \subset \mathbb{R}$ es conexo, ¿es $f^{-1}(K)$ necesariamente conexo?

b. Muéstrese que si $f: \mathbb{R}^n \to \mathbb{R}^m$ es continua en todo $\mathbb{R}^n$ y $B \subset \mathbb{R}$ está acotado, entonces $f(B)$ está acotado.

4. Analícese por qué es necesario imponer la condición $x \neq x_0$ en la definición de $\lim_{x\to x_0} f(x)$, considerando lo que podría ocurrir en el caso $f: \mathbb{R} \to \mathbb{R}$, $f(x) = 0$ si $x \neq 0$ y $f(0) = 1$ y cómo podríamos definir la derivada.

5. Muéstrese que $f: A \to \mathbb{R}^m$ es continua en x_0 sii para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $\|x - x_0\| \leq \delta$ implica $\|f(x) - f(x_0)\| \leq \varepsilon$. ¿Podemos reemplazar $\varepsilon > 0$ o $\delta > 0$ por $\varepsilon \geq 0$ o $\delta \geq 0$?

6. **a.** Sea $\{c_k\}$ una sucesión en $\mathbb{R}$. Muéstrese que $c_k \to c$ sii toda subsucesión de c_k tiene a su vez una subsucesión que converge a c.

b. Sea $f: \mathbb{R} \to \mathbb{R}$ una función acotada. Demuéstrese que f es continua si y sólo si la gráfica de f es un subconjunto cerrado de $\mathbb{R}^2$. ¿Qué ocurre si f no está acotada?

7. Considérese un conjunto compacto $B \subset \mathbb{R}^n$ y sea $f: B \to \mathbb{R}^m$ continua e inyectiva. Entonces, demuéstrese que $f^{-1}: f(B) \to B$ es continua. Muéstrese, mediante un ejemplo, que esto puede fallar si B es conexo, pero no compacto. (Para encontrar un contraejemplo, es necesario que $m > 1$.)

8. Definamos las aplicaciones $s: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ y $m: \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ como la suma y la multiplicación por un escalar, dadas por $s(x, y) = x + y$ y $m(\lambda, x) = \lambda x$. Muéstrese que estas aplicaciones son continuas.

9. Demuéstrese el siguiente "lema del empalme". *Sean $f: [a, b[\to \mathbb{R}^m$ y $g: [b, c] \to \mathbb{R}^m$ continuas. Definamos $h: [a, c] \to \mathbb{R}^m$ como $h = f$ en $[a, b[$ y $h = g$ en $[b, c]$. Si $f(b) = g(b)$, entonces h es continua.* Generalícese este resultado a conjuntos A, B en un espacio métrico.

10. Muéstrese que $f: A \to \mathbb{R}^m$, $A \subset \mathbb{R}^n$ es continua sii para todo conjunto $B \subset A$, $f(\text{cl}(B) \cap A) \subset \text{cl}(f(B))$.

11. **a.** Para $f:]a, b[\to \mathbb{R}$, muéstrese que si f es continua, entonces su gráfica Γ es conexa por arcos. Proporciónese un argumento intuitivo para el hecho de que si la gráfica de f es conexa por arcos, entonces f es continua. (Esto último es cierto, pero es un poco más difícil de demostrar.)

b. Sea $f: A \to \mathbb{R}^m, A \subset \mathbb{R}^n$; muéstrese que para $n \geq 2$, la conexión de la gráfica no implica la continuidad. [Sugerencia: para $f: \mathbb{R}^2 \to \mathbb{R}$, hágase una ranura en la gráfica.]

c. Analícese **b** para $m = n = 1$. [Sugerencia: en $\mathbb{R}$, considérese $f(x) = 0$ si $x = 0$ y $f(x) = \text{sen}(1/x)$ si $x > 0$.]

12. **a.** Una aplicación $f: A \subset \mathbb{R}^n \to \mathbb{R}^m$ se denomina ***Lipschitz en*** A si existe una constante $L \geq 0$ tal que $\|f(x) - f(y)\| \leq L\|x - y\|$ para todo $x, y \in A$. Muéstrese que una aplicación Lipschitz es uniformemente continua.

b. Encuéntrese una función continua acotada $f: \mathbb{R} \to \mathbb{R}$ que no sea uniformemente continua y, por lo tanto, no sea de Lipschitz.

c. ¿Es la suma (producto) de dos funciones Lipschitz también una función Lipschitz?

d. ¿Es la suma (producto) de dos funciones uniformemente continuas también una función uniformemente continua?

e. Sea f definida y con derivada continua en $]a - \varepsilon, b + \varepsilon[$, para algún $\varepsilon > 0$. Muéstrese que f es una función Lipschitz en $[a, b]$.

13. Sea f una función continua acotada $f: \mathbb{R}^n \to \mathbb{R}$. Demuéstrese que $f(U)$ es abierto para todo conjunto abierto $U \subset \mathbb{R}^n$ sii para todo conjunto abierto no vacío $V \subset \mathbb{R}^n$,

$$\inf_{x \in V} f(x) < f(y) < \sup_{x \in V} f(x)$$

para todo $y \in V$.

14. **a.** Encuéntrese una función $f: \mathbb{R}^2 \to \mathbb{R}$ tal que

$$\lim_{x \to 0} \lim_{y \to 0} f(x, y) \quad \text{y} \quad \lim_{y \to 0} \lim_{x \to 0} f(x, y)$$

existan pero que sean diferentes.

b. Encuéntrese una función $f: \mathbb{R}^2 \to \mathbb{R}$ tal que los dos límites en **a** existan y sean iguales pero que f no sea continua. [Sugerencia: $f(x, y) = xy/(x^2 + y^2)$ con $f = 0$ en $(0, 0)$.]

c. Encuéntrese una función $f: \mathbb{R}^2 \to \mathbb{R}$ que sea continua en cada recta que pase por el origen, pero que no sea continua. [Sugerencia: analícese la sugerencia de **b**, o considérese la función dada en coordenadas polares por $r \tan(\theta/4)$, $0 \le r < \infty$, $0 \le \theta < 2\pi$.]

15. Sean $f_1, \ldots, f_N$ funciones de $A \subset \mathbb{R}^n$ en $\mathbb{R}$. Sea m_i el supremo de f_i; es decir, $m_i = \sup(f_i(A))$. Sean $f = \sum f_i$ y $m = \sup(f(A))$. Muéstrese que $m \le \sum m_i$. Encuéntrese un ejemplo donde falle la igualdad.

16. Considérese una función $f: A \times B \to \mathbb{R}^m$, donde $A \subset \mathbb{R}$, $B \subset \mathbb{R}^r$. Decimos que f es ***parcialmente continua*** si para cada $x_0 \in A$ fijo, la aplicación $g(y) = f(x_0, y)$ es continua, y para cada $y_0 \in B$, $h(x) = f(x, y_0)$ es continua. Decimos que f es continua en A ***uniformemente con respecto a*** B si para cada $\varepsilon > 0$ y $x_0 \in A$, existe $\delta > 0$ tal que $\|x - x_0\| < \delta$ implica $\|f(x, y) - f(x_0, y)\| < \varepsilon$ para todo $y \in B$. Muéstrese que si f es parcialmente continua y es continua en A uniformemente con respecto de B, entonces f es continua.

17. Demuéstrese que las aplicaciones multilineales en el espacio euclídeo no son necesariamente uniformemente continuas.

18. Sean $A \subset M$ un conjunto conexo y $f: A \to \mathbb{R}$ una función continua, con $f(x) \neq 0$ para todo $x \in A$. Muéstrese que o bien $f(x) > 0$ para todo $x \in A$ o bien $f(x) < 0$ para todo $x \in A$.

19. Encuéntrese una aplicación continua $f: \mathbb{R}^n \to \mathbb{R}^m$ y un conjunto cerrado $A \subset \mathbb{R}^n$ tales que $f(A)$ no sea cerrado. De hecho, considérese el caso en que $f: \mathbb{R}^2 \to \mathbb{R}$ es la proyección sobre el eje x.

20. Proporciónese una demostración alternativa del teorema **4.4.1** por medio de la compacidad secuencial.

21. ¿Cuáles de las siguientes funciones en $\mathbb{R}$ son uniformemente continuas?

a. $f(x) = 1/(x^2 + 1)$.

b. $f(x) = \cos^3 x$.

c. $f(x) = x^2/(x^2 + 2)$.

d. $f(x) = x \operatorname{sen} x$.

22. Proporciónese una demostración alternativa del teorema de continuidad uniforme por medio del teorema de Bolzano-Weierstrass, de la manera siguiente. En primer lugar, muéstrese que si f no es uniformemente continua, existe $\varepsilon > 0$ y existen sucesiones x_n, y_n tales que $d(x_n, y_n) < 1/n$ y $d(f(x_n), f(y_n)) \geq \varepsilon$. Pásese a las subsucesiones convergentes y obténgase una contradicción con la continuidad de f.

23. Sea X un espacio métrico compacto y $f: X \to X$ una isometría; es decir, $d(f(x), f(y)) = d(x, y)$ para todo $x,y \in X$. Muéstrese que f es una biyección.

24. Sea $f: A \subset M \to N$.

a. Demuéstrese que f es uniformemente continua en A sii para cada par de sucesiones x_k, y_k de A tales que $d(x_k, y_k) \to 0$, tenemos que $d(f(x_k), f(y_k)) \to 0$.

b. Sea f uniformemente continua y sea x_k una sucesión de Cauchy en A. Muéstrese que $f(x_k)$ es una sucesión de Cauchy.

c. Sea f uniformemente continua y N completo. Muéstrese que f tiene una extensión única a una función continua en $\operatorname{cl}(A)$.

25. Sea $f:]0, 1] \to \mathbb{R}$ derivable y $f'(x)$ acotada. Muéstrese que $\lim_{x\to 0^+} f(x)$ y $\lim_{x\to 1^-} f(x)$ existen. Hágase esto en forma directa y aplicando el ejercicio 24c. Encuéntrese un contraejemplo si $f'(x)$ no está acotada.

26. Sea $f:]a, b] \to \mathbb{R}$ continua y tal que $f'(x)$ existe en $]a, b[$ y que $\lim_{x\to a^-} f'(x)$ exis-te. Demuéstrese que f es uniformemente continua.

27. Encuéntrese la suma de la serie $\sum_{k=4}^{\infty} (3/4)^k$.

28. Sea $f:]0, 1[\to \mathbb{R}$ uniformemente continua. ¿Debe f estar acotada?

29. Sea $f: \mathbb{R} \to \mathbb{R}$ tal que $|f(x) - f(y)| \leq |x - y^2|$. Demuéstrese que f es constante. [Sugerencia: ¿cuánto vale $f'(x)$?]

30. **a.** Sea $f: [0, \infty[\to \mathbb{R}$, $f(x) = \sqrt{x}$. Demuéstrese que f es uniformemente continua.

b. Sean $k > 0$ y $f(x) = (x - x^k)/\log x$ para $0 < x < 1$ y $f(0) = 0$, $f(1) = 1 - k$. Muéstrese que $f: [0, 1] \to \mathbb{R}$ es continua. ¿Es f uniformemente continua?

31. Sea $f(x) = x^{1/(x-1)}$ para $x \neq 1$. ¿Cómo podríamos definir $f(1)$ para que f sea continua en $x = 1$?

32. Sea $A \subset \mathbb{R}^n$ un conjunto abierto, $x_0 \in A$, $r_0 > 0$ y $B_{r_0} = \{x \in \mathbb{R}^n \mid \|x - x_0\| \leq r_0\}$. Supongamos que $B_{r_0} \subset A$. Demuéstrese que existe $r > r_0$ tal que $B_r \subset A$.

33. Un conjunto $A \subset \mathbb{R}^n$ es ***relativamente compacto*** si cl(A) es compacto. Demuéstrese que A es relativamente compacto sii toda sucesión en A tiene una subsucesión que converge a un punto en $\mathbb{R}^n$.

34. Supóngase que la temperatura en la superficie de la Tierra es una función continua y demuéstrese que en cualquier círculo máximo de ésta existen dos puntos antípodas con la misma temperatura.

35. Sea $f: \mathbb{R} \to \mathbb{R}$ creciente y acotada superiormente. Demuéstrese que el límite $\lim_{x \to 0^+} f(x)$ existe.

36. Muéstrese que $\{(x\ \mathrm{sen}(1/x)) \mid x > 0\} \cup (\{0\} \in [-1, 1])$ en $\mathbb{R}^2$ es conexo pero no conexo por arcos.

37. Demuéstrese el siguiente teorema de los valores intermedios para derivadas: si f es derivable en todo punto de $[a, b]$ y si $f'(a)$ y $f'(b)$ tienen signos contrarios, entonces existe un punto $x_0 \in\ ]a, b[$ tal que $f'(x_0) = 0$.

38. Una función de valores reales definida en $]a, b[$ es ***convexa*** si cumple la siguiente desigualdad para x, y en $]a, b[$ y para t en $[0, 1]$:

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y).$$

Si f tiene segunda derivada continua y $f'' > 0$, muéstrese que f es convexa.

39. Supóngase que f es continua en $[a, b]$, $f(a) = f(b) = 0$ y $x^2 f''(x) + 4xf'(x) + 2f(x) \geq 0$ para $x \in\ ]a, b[$. Demuéstrese que $f(x) \leq 0$ para x en $[a, b]$.

40. Calcúlese

$$\frac{d}{dt} \int_0^t \frac{dx}{1 + x^2}.$$

41. Demuéstrese que

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$

y

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

Estas fórmulas se utilizaron en el ejemplo resuelto **4.6.**

42. Para $x > 0$, defínase $L(x) = \int_1^x (1/t)dt$. Demuéstrese lo siguiente, usando esta definición:

- **a.** L es creciente en x.
- **b.** $L(xy) = L(x) + L(y)$.
- **c.** $L'(x) = 1/x$.
- **d.** $L(1) = 0$.
- **e.** Las propiedades **c** y **d** determinan L de manera única. ¿Qué es L?

43. Sea $f: \mathbb{R} \to \mathbb{R}$ continua y definamos $F(x) = \int_0^{x^2} f(y)dy$. Demuéstrese que $F'(x) = 2xf(x^2)$. Proporciónese un teorema más general.

44. Sea $f: [0, 1] \to \mathbb{R}$ integrable Riemann y supóngase que para cada a, b tales que $0 \leq a < b \leq 1$ existe c, $a < b < c$ tal que $f(c) = 0$. Demuéstrese que $\int_0^1 f = 0$. ¿Debe anularse f? ¿Qué ocurre si f es continua?

45. Demuéstrese el siguiente ***segundo teorema del valor medio.*** *Sean f y g funciones definidas en $[a, b]$, con g continua, $f \geq 0$ y f integrable. Entonces existe un punto $x_0 \in \,]a, b[$ tal que*

$$\int_a^b f(x)g(x)\,dx = g(x_0)\int_a^b f(x)\,dx.$$

46. **a.** Para las funciones de valores complejos definidas en un intervalo, demuéstrese el teorema fundamental del cálculo.

b. Evalúese $\int_0^\pi e^{ix}\,dx$ usando **a.**

Capítulo 5
Convergencia uniforme

Muchas funciones importantes se definen mediante sucesiones o series infinitas. Para el estudio de tales funciones, necesitamos comprender el concepto de convergencia uniforme. Con el fin de tratar eficazmente situaciones y ejemplos concretos echaremos mano de criterios específicos para la convergencia uniforme; de éstos probablemente el más útil sea el criterio M de Weierstrass para series. Otro criterio es el de Cauchy, que se emplea principalmente con fines teóricos. También incluiremos los criterios más particulares de Dirichlet y Abel.

En conexión con la convergencia uniforme volveremos a tratar el espacio de las funciones continuas, que introdujimos brevemente en §1.7. En este espacio, los "puntos" o "vectores" son funciones, y la convergencia de una sucesión corresponde a la convergencia uniforme de estas funciones. Demostraremos que este espacio es completo, en el sentido de que las sucesiones de Cauchy convergen. Una segunda propiedad básica de este espacio, el teorema de Arzela-Ascoli, establecerá la compacidad de ciertos subconjuntos. Otro resultado importante, el teorema de Stone-Weierstrass, permitirá aproximar las funciones continuas por polinomios o por funciones de otras clases adecuadas. En relación con una técnica iterativa básica llamada el principio de la aplicación contractiva, daremos algunas aplicaciones a las ecuaciones diferenciales e integrales y algunos problemas de la teoría de control.

§5.1 Convergencias uniforme y puntual

Hay varias formas distintas de concebir la convergencia de una sucesión de funciones, de las cuales la más natural sería la *convergencia puntual*. Con esta idea, sólo exigimos que para cada punto x del dominio, la sucesión de valores $f_k(x)$ converja.

5.1.1 Definición *Sean N un espacio métrico, A un conjunto y $f_k: A \to N$, $k = 1, 2, \ldots$. Esta sucesión de funciones f_k* ***converge puntualmente (o simplemente)*** *a una función $f: A \to N$ si para cada $x \in A$, $f_k(x) \to f(x)$ (la convergencia como una sucesión en N). Con frecuencia escribimos $f_k \to f$ (puntualmente) si f_k converge puntualmente a f.*

Aunque este tipo de convergencia es útil para ciertos fines, hay otros casos en que no lo es. Si uno calcula numéricamente f aproximándola mediante las funciones f_k, no está claro que la convergencia puntual sea el concepto adecuado; he aquí algo que puede fallar: aun cuando las funciones f_k sean continuas, f no tiene por qué serlo. Por ejemplo, consideremos la figura 5.1-1, en la cual

$$f_k(x) = \begin{cases} 0, & x \geq 1/k, \\ -kx + 1, & 0 \leq x < 1/k. \end{cases}$$

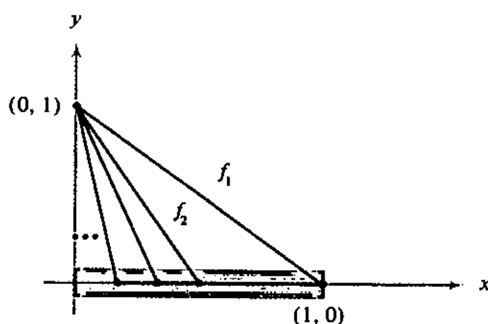


FIGURA 5.1-1 Convergencia puntual

En este caso, la sucesión $f_k(x)$ converge para cada x en el intervalo $[0, 1]$. La sucesión $f_k(x)$ converge a 0 si $x > 0$ y a 1 si $x = 0$. Así, la función límite queda definida por $f(x) = 0$ si $x > 0$ y $f(0) = 1$. Esta función no es continua, aunque cada f_k lo sea.

¿Qué está fallando? Si nos fijamos en cualquier valor de x, entonces, para k suficientemente grande, los valores $f_k(x)$ son cercanos a $f(x)$; sin embargo, el valor necesario de k para cualquier grado particular de precisión depende mucho de x. Para x cercano a 0, se necesitarán valores grandes de k para que $f_k(x)$ esté cerca de $f(x)$ y siempre habrá puntos x más cercanos a 0 para los cuales $f_k(x)$ no esté cerca de $f(x)$. Para remediar esta situación, vamos a introducir un concepto más fuerte de convergencia. Exigiremos no sólo que los valores de f_k estén cerca de los de f, sino que, además, lo estén uniformemente; es decir, independientemente de x.

5.1.2 Definición *Sea $f_k : A \to N$ una sucesión de funciones con la propiedad de que para cada $\varepsilon > 0$ existe un entero L tal que $k \geq L$ implica que $\rho(f_k(x), f(x)) < \varepsilon$ para todo $x \in A$. Aquí, ρ es la métrica en N. Bajo estas condiciones, decimos que la sucesión f_k* ***converge uniformemente a f en A****, y escribimos $f_k \to f$ (uniformemente).*

En esta definición, el punto importante es que la elección de N puede depender de ε pero *no* de x; debemos poder elegir N de modo que el mismo valor de N sirva para todo el conjunto A. Así, el concepto de convergencia uniforme depende en gran medida del conjunto A considerado como dominio. La convergencia puede ser uniforme en un dominio, pero podría no serlo en un conjunto más grande. En el último ejemplo, la convergencia de las funciones f_k al límite f es uniforme en cualquier intervalo $[a, 1]$, con $0 < a < 1$. Dado $\varepsilon > 0$, sólo debemos elegir N entero tal que $1/N < a$. Si $k \geq N$, entonces $f_k(x) = 0 = f(x)$ para cada x en $[a, 1]$. Así, $f_k \to f$ uniformemente en $[a, 1]$. Sin embargo, si a es muy próximo a 0, entonces se necesitan valores correspondientes de N cada vez mayores. Si $\varepsilon < 1/2$, entonces es imposible elegir un valor de k tal que $|f_k(x) - f(x)| < \varepsilon$ para todo x en $[0, 1]$ al mismo tiempo. Así, aunque las funciones f_k convergen a f puntualmente en $[0, 1]$, no lo hacen uniformemente.

Hay una buena forma geométrica de visualizar la convergencia uniforme. La condición $\rho(f_k(x), f(x)) < \varepsilon$ para todo x significa que $f_k(x)$ siempre está más cerca de $f(x)$ que una distancia ε. Esto puede describirse más fácilmente en términos de las gráficas de las funciones. La gráfica de f_k debe estar dentro de un "tubo de anchura ε" alrededor de la gráfica de f, si $k \geq N$. Véase la figura 5.1-2.

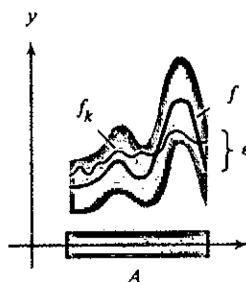


FIGURA 5.1-2 Proximidad uniforme para $f : A \subset \mathbb{R} \to \mathbb{R}$

Es probable que otro ejemplo aclare este concepto. Considérese la sucesión de funciones $f_k : \mathbb{R} \to \mathbb{R}$ definidas por

$$f_k(x) = \begin{cases} 0, & x < k, \\ 1, & x \geq k \end{cases}$$

($k = 1, 2, 3, \ldots$). Entonces $f_k \to 0$ (puntualmente), pues para cada $x \in \mathbb{R}$, $f_k(x) = 0$ cuando k es grande ($k > x$). Sin embargo, f_k no converge a cero uniformemente, ya que independientemente de lo grande que sea k, existen puntos x para los que $f_k(x) - 0$ no es pequeño.

Obsérvese que si $f_k(x) \to f$ (uniformemente), entonces $f_k \to f$ (puntualmente). Esto se debe a que para cada $x \in A$ y $\varepsilon > 0$, tenemos un entero L tal que $\rho(f_k(x), f(x)) < \varepsilon$ si $k \geq L$, es decir, $f_k(x) \to f(x)$.

Se pueden hacer definiciones similares para una serie de funciones. En este caso, elegiremos $N = V$, un espacio vectorial normado y definiremos $g_k : A \to V$, con el fin de que la suma tenga sentido.

5.1.3 Definición *Decimos que la serie $\sum_{k=1}^{\infty} g_k$ converge a g **puntualmente,** y escribimos $\sum_{k=1}^{\infty} g_k = g$ (puntualmente), si la sucesión $s_k = \sum_{i=1}^{k} g_i$ de sumas parciales converge puntualmente a g. Además, decimos que $\sum_{k=1}^{\infty} g_k = g$ (uniformemente) o que $\sum_{k=1}^{\infty} g_k$ converge a g **uniformemente** si $s_k \to g$ (uniformemente). Para una sucesión f_k (o serie $\sum g_k$), decimos que f_k (o $\sum g_k$) **converge uniformemente** si existe una función a la que converge uniformemente.*

La primera propiedad básica de la convergencia uniforme es que preserva la continuidad, como muestra el siguiente resultado.

5.1.4 Proposición *Sean $A \subset M$, donde M es un espacio métrico, y $f_k : A \to N$ una sucesión de funciones continuas tales que $f_k \to f$ (uniformemente en A). Entonces f es continua en A.*

Así, la convergencia uniforme es una condición lo bastante fuerte como para garantizar que la función límite de una sucesión de funciones continuas sea continua. En vista de los ejemplos anteriores, esto debería parecer razonable.

5.1.5 Corolario *Si las funciones $g_k : A \to V$ son continuas y $\sum_{k=1}^{\infty} g_k = g$ (uniformemente), entonces g es continua.*

Éste se sigue de aplicar la proposición a la sucesión de las sumas parciales. El resultado se puede enunciar también como la validez del ***intercambio de límites y sumatorias***:

$$\lim_{x \to x_0} \sum_{k=1}^{\infty} g_k(x) = \sum_{k=1}^{\infty} \lim_{x \to x_0} g_k(x)$$

5.1.6 Ejemplo *Sean $f_n : \mathbb{R} \to \mathbb{R}$ dadas por $f_n(x) = (\operatorname{sen} x)/n$. Muéstrese que $f_n \to 0$ uniformemente cuando $n \to \infty$.*

Solución Debemos probar que $|f_n(x) - 0| = |f_n(x)|$ se hace pequeño independientemente de x cuando $n \to \infty$. Pero $|f_n(x)| = |\text{sen}\, x|/n \leq 1/n$, que, efectivamente, se hace pequeño independientemente de x cuando $n \to \infty$. ♦

5.1.7 Ejemplo *Muéstrese que la serie de* sen x,

$$\text{sen}\, x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots,$$

converge uniformemente en el intervalo $[0, r]$ *para todo* $r > 0$.

Solución Debemos probar que la sucesión de sumas parciales

$$s_n(x) = \sum_{k=0}^{n} \frac{(-1)^k x^{2k+1}}{(2k+1)!}$$

converge uniformemente. Para ello, estimaremos la diferencia

$$|s_n(x) - \text{sen}\, x| = \left| \sum_{k=n+1}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \right| \leq \sum_{k=n+1}^{\infty} \frac{r^{2k+1}}{(2k+1)!}.$$

El miembro derecho es independiente de x y tiende a 0 cuando n tiende a ∞, pues es la cola de una serie convergente. Así pues, la convergencia de esta serie es uniforme. Obsérvese que la continuidad de sen x se sigue de este hecho, resultado que ya conocíamos anteriormente. ♦

5.1.8 Ejemplo *Sean* $f_n(x) = x^n$, $0 \leq x \leq 1$. *¿Converge* f_n *uniformemente? ¿Qué ocurre si* $0 \leq x < 1$?

Solución Determinemos primero el límite puntual. Tenemos que $f_n(0) = 0$ para todo n y $f_n(x) \to 0$ si $x < 1$, pero $f_n(1) = 1$ para todo n. Así que f_n converge puntualmente a

$$f(x) = \begin{cases} 0, & x \neq 1 \\ 1, & x = 1. \end{cases}$$

No puede converger uniformemente, ya que este límite no es continuo (figura 5.1-3). Si $0 \leq x < 1$, entonces f_n converge puntualmente a cero y 0 es una función continua. Sin embargo, la convergencia tampoco es uniforme, ya que dado cualquier n, $x^n \geq 1/2$ si x está suficientemente cerca de 1 (ya que $\lim_{x \to 1} x^n = 1$). ♦

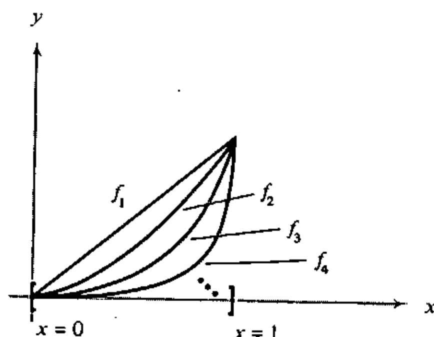


FIGURA 5.1-3 La sucesión $f_n(x) = x^n$

5.1.9 Ejemplo *Considérese la serie geométrica* $\sum_{k=0}^{\infty} x^k$.

a. *Muéstrese que esta serie converge puntualmente a la función* $g(x) = 1/(1-x)$ *para* x *en el intervalo abierto* $]-1, 1[$.

b. *Si* $0 < a < 1$, *muéstrese que la convergencia es uniforme en el intervalo* $[-a, a]$.

c. *Muéstrese que la convergencia no es uniforme en* $]-1, 1[$.

Solución

a. Las sumas parciales están dadas por

$$s_n(x) = \sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1-x},$$

como se puede verificar al multiplicar $(1-x)(1+x+x^2+\cdots+x^n$. Así, $|s_n(x) - g(x)| = |x|^{n+1}/|1-x|$. Para x fija en el intervalo abierto $]-1, 1[$, el denominador $|1-x|$, está fijo, mientras que el numerador, $|x|^{n+1}$, tiende a 0 cuando n tiende a ∞. Más precisamente, podemos calcular que $|x|^{n+1}/|1-x| < \varepsilon$, siempre que $n > (\varepsilon|1-x|/\log(x)) - 1$.

b. El error $|x|^{n+1}/|1-x| < \varepsilon$ es menor para $-|x|$ que para $|x|$, y

$$\frac{d}{dx}\left(\frac{|x|^{n+1}}{1-x}\right) = \left(\frac{nx^n(1-x)+x^n}{(1-x)^2}\right)$$

(¿por qué?). Como esta derivada es positiva, vemos que el error en $[-a, a]$ es mayor en el extremo derecho, $x = a$. Si seleccionamos N lo bastante grande como

para que sirva para ese punto, también servirá en todo $[-a, a]$ y se cumplirá la definición de convergencia uniforme en $[-a, a]$.

c. Si $x > 1/2$, entonces el error, $|x^{n+1}/(1-x^n)|$, será mayor que $2|x^{n+1}|$. Así, para cada valor particular de n, podemos hacer que el error sea mayor que cualquier ε dado entre 0 y 2 eligiendo x en el intervalo abierto $]1/2, 1[$ y mayor que $(\varepsilon/2)^{1/(n+1)}$. Así, la convergencia no es uniforme en todo $]-1, 1[$. ♦

5.1.10 Ejemplo *Sean* $f_n(x) = x^n/(1 + x^n)$ *para* x *en el intervalo* $[0, 2]$. *Muéstrese que la sucesión de funciones* $f_1, f_2, f_3, \ldots$ *converge puntualmente en el intervalo* $[0, 2]$ *pero que la convergencia no es uniforme.*

Solución El denominador $1 + x^n$ es mayor o igual que 1 para cada x del dominio, por lo que $|f_n(x)| \le x^n$. Si $0 \le x < 1$, esto tiende a 0 cuando $n \to \infty$. Por lo tanto, $\lim_{n\to\infty} f_n(x) = 0$ si $0 \le x < 1$. Si $x = 1$, entonces $f_n(x) = 1/2$ para cada n. Si $x > 1$, entonces $x^n \to \infty$ cuando $n \to \infty$. Como $f_n(x) = 1/(1 + (1/x^n))$, esto implica que $\lim_{n\to\infty} f_n(x) = 1$ para $x > 1$. Así pues, f_n tiende puntualmente en $[0, 2]$ a la función $f(x)$ dada por $f(x) = 0$ para $0 \le x < 1$, $f(1) = 1/2$ y $f(x) = 1$ para $1 < x \le 2$. Cada una de las funciones f_n es continua en $[0, 2]$. Si la convergencia fuera uniforme, entonces la función límite sería continua. Como f tiene una discontinuidad en $x = 1$, la convergencia no puede ser uniforme.

Hemos dibujado las primeras funciones en la figura 5.1-4, junto con la función límite discontinua. ♦

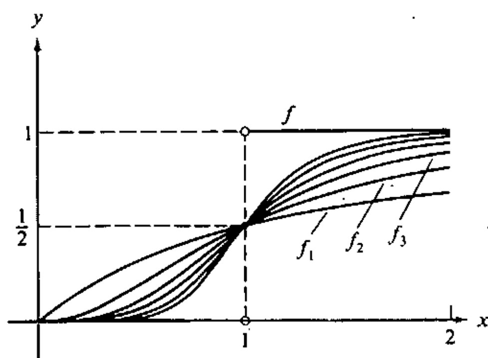


FIGURA 5.1-4 Esta sucesión f_k converge puntualmente pero no uniformemente

Ejercicios de §5.1

1. Sea $f_n(x) = (x - 1/n)^2$, $0 \leq x \leq 1$. ¿Converge f_n uniformemente?

2. Sea $f_n(x) = x - x^n$, $0 \leq x \leq 1$. ¿Converge f_n uniformemente?

3. Sean $f_n : \mathbb{R} \to \mathbb{R}$ uniformemente continuas tales que f_n converge uniformemente a f. ¿Piensa el lector que f debe ser uniformemente continua? Analícese.

4. Sea $f_n(x) = x^n$, $0 \leq x \leq 0.999$. ¿Converge f_n uniformemente?

5. Sea

$$f(x) = \sum_{n=1}^{\infty} \frac{x^{n/2}}{n(n!)^2} \quad 0 \leq x \leq 1.$$

Analícese cómo podría demostrarse que f es continua.

§5.2 Criterio M de Weierstrass

Una razón por la que hay pocos ejemplos de series infinitas en la sección anterior es que, por lo general, es difícil estudiar la convergencia uniforme directamente, a menos que se disponga de una fórmula conveniente para las sumas parciales, como es el caso de la serie geométrica. En esta sección vamos a introducir algunos criterios convenientes para la convergencia uniforme de una serie de funciones. El primero es una caracterización de aquellas series de funciones que son uniformemente convergentes expresada en términos de las funciones de la serie, sin hacer referencia explícita a la función límite a la cual converge la serie. La idea básica es usar la condición de Cauchy para la convergencia de la sucesión de sumas parciales. Recuérdese que toda sucesión convergente satisface la condición de Cauchy. Éstas son las sucesiones que, en cierto sentido, "deben" converger. Si el espacio en que están los puntos es completo, entonces el límite realmente existe y la sucesión converge.

5.2.1 Criterio de Cauchy *Sea N un espacio métrico con métrica ρ y sea A un conjunto. Supóngase que N es completo y que $f_k : A \to N$ es una sucesión de funciones. Entonces f_k converge uniformemente en A sii para cada $\varepsilon > 0$ existe N tal que $l, k \geq N$ implica*

$$\rho(f_k(x), f_l(x)) < \varepsilon$$

para todo $x \in A$.

Para una serie, la condición de Cauchy adquiere la siguiente forma, al aplicarla a la sucesión de sumas parciales. *La serie* $\sum_{k=1}^{\infty} g_k$ *converge uniformemente en A sii para cada* $\varepsilon > 0$ *existe N tal que* $k \geq N$ *implica*

$$\|g_k(x) + \cdots + g_{k+p}(x)\| < \varepsilon$$

para cada $x \in A$ *y para todo entero* $p \geq 0$. Como hicimos en **§5.1** para series, supondremos que g_k toma valores en un espacio vectorial normado V. Para usar **5.2.1**, necesitamos que V sea *completo*. Por supuesto, $V = \mathbb{R}^n$ e incluso $V = \mathbb{R}$ son los casos más frecuentes.

Por medio de la condición de Cauchy podemos obtener la siguiente técnica importante para determinar la convergencia uniforme de una serie.

5.2.2 Criterio *M* de Weierstrass *Supongamos que* V *es un espacio vectorial normado completo y que* $g_k : A \to V$ *son funciones tales que existen constantes* M_k *que cumplen* $\|g_k(x)\| \leq M_k$ *para cada* $x \in A$ *y* $\sum_{k=1}^{\infty} M_k$ *converge. Entonces* $\sum_{k=1}^{\infty} g_k$ *converge uniformemente (y absolutamente).*

No siempre es posible usar el criterio M, pero es muy eficaz en la mayoría de los casos. Si se quieren criterios más elaborados, véanse los de Dirichlet y Abel en **§5.9**.

En el criterio M, las constantes M_k proporcionan una cota de la "tasa de convergencia", y la clave está en que tal cota no depende de x. Más exactamente, la cola de la serie $\sum_{k=1}^{\infty} g_k$, que representa el error, está acotada por la cola de $\sum_{k=1}^{\infty} M_k$, que tiende a cero independientemente de x.

5.2.3 Ejemplo *Muéstrese que la serie* $\sum_{n=1}^{\infty} g_n(x) = \sum_{n=1}^{\infty} (\text{sen } nx)^2/n^2$ *converge uniformemente en* $\mathbb{R}$.

Solución Sean $M_n = 1/n^2$. Entonces $|g_n(x)| \leq M_n$, ya que $|\text{sen } nx| \leq 1$. También sabemos que $\sum(1/n^2)$ converge, pues es una serie de Riemann con $p = 2 > 1$. Por lo tanto, por el teorema **5.2.2**, la serie converge uniformemente. ♦

5.2.4 Ejemplo *Demuéstrese que* $f(x) = \sum_{n=0}^{\infty} \left(\frac{x^n}{n!}\right)^2$ *es continua en* $\mathbb{R}$.

Solución En este caso no podemos elegir M_n para el n-ésimo término, porque x no está acotada. No esperamos, por lo tanto, que se cumpla la convergencia uniforme en todo $\mathbb{R}$, pero sí que podemos demostrar la convergencia uniforme en cada intervalo

$[-a, a]$, tomando $M_n = (a^n/n!)^2$, que es una cota superior para el n-ésimo término en $[-a, a]$. El criterio del cociente muestra que $\sum M_n$ converge, pues

$$\frac{M_{n+1}}{M_n} = \left(\frac{n!}{(n+1)!}\right)^2 \left(\frac{a^{n+1}}{a^n}\right)^2 = \left(\frac{a}{n+1}\right)^2$$

converge a cero, que es menor que 1. Por lo tanto, tenemos convergencia uniforme en $[-a, a]$ y, por **5.1.5**, podemos deducir la continuidad de f en $[-a, a]$. Como a es arbitrario, obtenemos la continuidad en todo $\mathbb{R}$. ♦

5.2.5 Ejemplo *Supongamos que una sucesión $f_n(x)$, $0 \leq x \leq 1$, converge uniformemente y que f_n es derivable. ¿Debe $f_n'(x)$ converger uniformemente?*

Solución La respuesta es no. Por lo general, el control sobre las derivadas proporciona el control sobre las funciones, por medio del teorema del valor medio, pero no viceversa. Por ejemplo, sea $f_n(x) = [\text{sen}(n^2x)]/n$. Entonces, $f_n \to 0$ uniformemente, pero $f_n'(x) = n\cos(n^2x)$ no converge ni siquiera puntualmente (por ejemplo, sea $x = 0$). He aquí otro ejemplo de este fenómeno. Sea $g_n(x) = x^{n+1}/(n+1)$ en $[0, 1]$. Como $|g_n(x)| \leq 1/(n+1)$, $g_n \to 0$ uniformemente, pero $g_n'(x) = x^n$ *no* converge uniformemente, como ya demostramos. ♦

5.2.6 Ejemplo *Supongamos que $a_0, a_1, a_2, \ldots$ es una sucesión acotada de números reales. Muéstrese que la serie*

$$\sum_{k=0}^{\infty} \frac{a_k}{k!} x^k$$

converge a una función continua.

Solución Fijemos un punto x_0. Necesitamos demostrar la convergencia a una función límite $f(x_0)$ y la continuidad de f en x_0. Sea $A = \{x \mid |x| \leq 2|x_0|\}$. Entonces $x_0 \in A$. Las sumas parciales son polinomios y ciertamente son continuas en el conjunto A. Si probásemos que la serie converge uniformemente en A, la suma sería continua en el conjunto A y ciertamente lo sería en x_0. Como x_0 es un punto arbitrario, esto completará la demostración. La convergencia uniforme en A se puede establecer por medio del criterio M de Weierstrass. Si B es una cota superior para los números $|a_k|$, podemos tomar $M_k = B \cdot 2^k|x_0|^k/k!$. Como $g_k(x) = (a_k/k!)x^k$, tenemos que $|g_k(x)| \leq M_k$ para todo x en A. El criterio del cociente muestra que $\sum_{k=0}^{\infty} M_k$ converge, pues $M_{k+1}/M_k = 2|x_0|/(k+1)$, que tiende a 0 cuando k tiende a ∞. El criterio de Weierstrass se aplica, pues, y prueba que nuestra serie converge uniformemente en A, como queríamos. ♦

Ejercicios de §5.2

1. Analícese la convergencia y la convergencia uniforme de

   a. $f_n(x) = x^n/(n + x^n)$, $x \geq 0$, $n = 1, 2, \ldots$.

   b. $f_n(x) = e^{-x^2/n}/n$, $x \in \mathbb{R}$, $n = 1, 2, \ldots$.

2. Analícese la convergencia uniforme de $\sum_{n=1}^{\infty} x^n/n^2$, $0 \leq x \leq 1$.

3. Demuéstrese que $f(x) = \sum_{n=1}^{\infty} x^n/n^2$ es continua en $[0, 1]$.

4. Analícese la convergencia uniforme de $\sum_{n=1}^{\infty} 1/(x^2 + n^2)$.

5. Si $\sum_{n=1}^{\infty} a_n$ es absolutamente convergente, demuéstrese que $\sum_{n=1}^{\infty} a_n \operatorname{sen} nx$ es uniformemente convergente.

§5.3 Integración y derivación de series

Hemos visto en las dos secciones anteriores que a menudo las series infinitas representan funciones continuas. Si queremos saber cómo se pueden usar en el estudio del cálculo, necesitamos conocer cómo se comportan bajo la derivación y la integración. Si una sucesión de funciones $f_1, f_2, f_3, \ldots$ converge a una función límite f, ¿bajo qué condiciones la sucesión $f_1', f_2', f_3', \ldots$ converge a f' o las integrales $\int f_1, \int f_2, \int f_3, \ldots$ convergen a $\int f$? La idea de convergencia uniforme proporciona una herramienta que con frecuencia resulta útil. El resultado relativo a integrales es casi directo.

5.3.1 Teorema *Supongamos que $f_1, f_2, f_3, \ldots$ son funciones integrables en un intervalo cerrado y acotado $[a, b]$ que convergen uniformemente a una función límite f en $[a, b]$. Entonces f es integrable en $[a, b]$* y

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx.$$

La idea es que, para n grande, los valores $f_n(x)$ están cerca de $f(x)$ de manera uniforme. Así pues, cualquier suma de Riemann de f estará próxima a la correspondiente suma de Riemann de f_n. La longitud finita del intervalo es importante en este paso. Como f_n es integrable, sus sumas de Riemann se aproximan entre sí cuando el tamaño de las particiones es lo bastante pequeño y, por tanto, los de f también se aproximan entre sí. En particular, las sumas inferior y superior de f se acercan entre sí. Con los

detalles adecuados, esto demuestra que f es integrable. La afirmación de que el límite de las integrales es igual a la integral del límite se sigue entonces de la siguiente observación:

$$\left| \int_a^b f_n(x)\,dx - \int_a^b f(x)\,dx \right| \leq \int_a^b |f_n(x) - f(x)|\,dx.$$

Si elegimos n lo bastante grande como para que $|f_n(x) - f(x)| < \varepsilon$ para todo $x \in [a, b]$, el miembro derecho es menor que $(b-a)\varepsilon$ y se obtiene, como resultado, nuestra afirmación sobre los límites. De nuevo, la longitud finita del intervalo es importante. Los detalles de la demostración aparecen al final del capítulo.

Si aplicamos este resultado a las sumas parciales de una serie infinita de funciones integrables que convergen uniformemente en un intervalo cerrado y acotado $[a, b]$, obtendremos que podemos intercambiar el orden de la integración y la suma.

5.3.2 Corolario *Supongamos que las funciones $g_k : [a, b] \to \mathbb{R}$ son integrables Riemann y que $\sum_{k=1}^{\infty} g_k$ converge uniformemente en $[a, b]$. Entonces podemos intercambiar el orden de la integración y la suma:*

$$\int_a^b \left(\sum_{k=1}^{\infty} g_k(x) \right) dx = \sum_{k=1}^{\infty} \left(\int_a^b g_k(x)\,dx \right)$$

El corolario es consecuencia del teorema **5.3.1**, aplicado a la sucesión de las sumas parciales.

Intuitivamente, el teorema debería ser claro, pues si f_k está muy cerca de f, entonces su integral (el área debajo de la curva) debe estar cerca de la de f. Pero hay que tener cuidado. De hecho, ¡este resultado puede ser falso si f_k sólo converge puntualmente! (Véase el ejemplo **5.3.5**.)

Nota. Existe un teorema de más amplia validez que el **5.3.1**, llamado el *teorema de la convergencia dominada de Lebesgue*. Una versión de este resultado (debida a Ascoli)[1] dice que si f_k converge puntualmente a f y si las f_k están uniformemente acotadas (es decir, $|f_k(x)| \leq M$ para todo $k = 1, 2, \ldots$ y todo $x \in [a, b]$), entonces la conclusión del teorema **5.3.1** sigue siendo válida. En la mayor parte de este libro nos conformaremos con la forma más elemental de este resultado que aparece en el teorema **5.3.1**.

¿Podemos tomarnos la misma libertad con las derivadas? Con frecuencia, la respuesta a la pregunta de la derivación término a término de una sucesión o serie uniformemente convergente es no, como vimos en el ejemplo **5.2.5**. Este resultado es un buen

[1] El resultado se debe a Arzelà y no a Ascoli. La referencia es: Luxemberg, W.A.J., "Arzelà's Dominated Convergence Theorem for the Riemann Integral", *Amer. Math. Monthly 78*, 970-979, 1971. (*N. del R.T.*)

ejemplo del tipo de cuidado que a menudo es necesario para transformar una afirmación intuitivamente plausible en un hecho real. Así pues, necesitamos más hipótesis además de la convergencia uniforme. En el teorema siguiente damos unas condiciones suficientes.

5.3.3 Teorema *Sea $f_k :]a, b[\to \mathbb{R}$ una sucesión de funciones derivables en el intervalo abierto $]a, b[$ que convergen puntualmente a $f :]a, b[\to \mathbb{R}$. Supongamos que las derivadas f'_k son continuas y que convergen uniformemente a una función g. Entonces f es derivable y $f' = g$.*

5.3.4 Corolario *Si las funciones g_k son derivables, las g'_k son continuas, $\sum_{k=1}^{\infty} g_k$ converge puntualmente y $\sum_{k=1}^{\infty} g'_k$ converge uniformemente, entonces*

$$\left(\sum_{k=1}^{\infty} g_k\right)' = \sum_{k=1}^{\infty} g'_k.$$

Como es habitual, el corolario es consecuencia de la aplicación del teorema a la sucesión de las sumas parciales.

5.3.5 Ejemplo *Proporciónese un ejemplo de una sucesión $f_k : [0, 1] \to \mathbb{R}$ que converja a cero puntualmente, pero para la cual $\int_a^b f_k \, dx$ no converja a cero.*

Solución Sea f_k la función con la gráfica de la figura 5.3-1. Entonces, f_k es tal que $\int_0^1 f_k \, dx = 1$ para todo $k = 1, 2, 3, \ldots$. Además, para cada x, $f_k(x) \to 0$ cuando $k \to \infty$ (claramente si $x = 0$, y si $x > 0$, entonces $f_k(x) = 0$ en el momento en que $k > 1/x$). Así pues, $f_k \to 0$ puntualmente, pero $\int_0^1 f_k \, dx = 1$ para todo k. ♦

5.3.6 Ejemplo *Sean $g_n(x) = nx^2/(1 + nx^2)$, $-1 \le x \le 1$, $n = 1, 2, 3, \ldots$. Analícense las conclusiones del teorema **5.3.3** en este caso.*

Solución Cuando n crece, las cantidades $f_n(x)$ tienden puntualmente, pero no uniformemente, a la función $f(x)$ definida en $[-1, 1]$ por $f(x) = 1$ si $x \neq 0$ y $f(0) = 0$. La función límite no es ni siquiera continua, mucho menos derivable, en $x = 0$. Algo debe fallar con las derivadas, las cuales están dadas por $f'_n(x) = 2nx/(1 + nx^2)^2$. Éstas convergen puntualmente a 0 en $[-1, 1]$, pero *no uniformemente*: para $n \ge 1$ obtenemos $f'_n(1/n) \ge 1/2$. Este

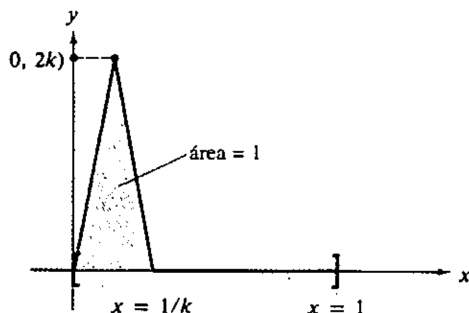


FIGURA 5.3-1 Esta sucesión converge puntualmente, pero la sucesión correspondiente de integrales no converge

ejemplo muestra que una hipótesis como la convergencia uniforme de las derivadas es importante en el teorema **5.3.3**; la convergencia puntual no es suficiente. ♦

5.3.7 Ejemplo *Verifíquese que* $\int_0^x e^t\,dt = e^x - 1$, *usando* $e^x = \sum_0^\infty x^n/n!$ *y el corolario* **5.3.2.**

Solución Por el criterio M de Weierstrass, $e^x = \sum_0^\infty x^n/n!$ converge uniformemente en cualquier intervalo finito. Así, por el corolario **5.3.2** aplicado al intervalo $[0, x]$,

$$\int_0^x e^t\,dt = \sum_{n=0}^{\infty} \int_0^x \frac{t^n}{n!}\,dt = \sum_{n=0}^{\infty} \frac{t^{n+1}}{(n+1)!}\bigg|_0^x = \frac{x}{1!} + \frac{x^2}{2!} + \cdots$$

$$= e^x - 1. \quad ♦$$

5.3.8 Ejemplo

a. *Súmese la serie* $\sum_{n=1}^{\infty} x^n/n$, $|x| < 1$.

b. *Súmese* $1 - 1/2 + 1/3 - 1/4 + \cdots$.

Solución

a. Sabemos que $\sum_{n=0}^{\infty} x^n = 1/(1-x)$, pues es una serie geométrica. La convergencia es uniforme en $[-1+\varepsilon, 1-\varepsilon]$ para todo $\varepsilon > 0$, por el criterio M. Así, podemos

integrar término a término de 0 a x:

$$\sum_{n=0}^{\infty} \frac{x^{n+1}}{n+1} = -\log(1-x); \quad \text{es decir,} \quad \sum_{n=1}^{\infty} \frac{x^n}{n} = -\log(1-x).$$

Esto es válido puntualmente para todo x en $]-1, 1[$, pues ε es arbitrario.

b. En realidad, es válido también en $x = -1$. Para verlo, recordemos que

$$\sum_{n=0}^{N} x^n = \frac{1-x^{N+1}}{1-x} = \frac{1}{1-x} - \frac{x^{N+1}}{1-x}.$$

Entonces,

$$\sum_{n=0}^{N} \int_0^x x^n\, dx = \sum_{n=0}^{N} \frac{x^{n+1}}{n+1} = -\log(1-x) - \int_0^x \frac{x^{N+1}}{1-x}\, dx$$

y, por lo tanto,

$$\left| \sum_{n=0}^{N} \frac{x^{n+1}}{n+1} + \log(1-x) \right| = \left| \int_0^x \frac{t^{N+1}}{1-t}\, dt \right|.$$

Si hacemos $x = -1$, usando $1/(1-t) \le 1$ para $-1 \le t < 0$, obtenemos

$$\left| \sum_{n=0}^{N} \frac{(-1)^{n+1}}{n+1} + \log 2 \right| \le \left| \int_0^{-1} t^{N+1}\, dt \right| = \frac{1}{N+2}$$

que tiende a 0 cuando N tiende a ∞. Así pues,

$$\log 2 = -\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots. \quad \blacklozenge$$

5.3.9 Ejemplo

a. *Desarróllese* $f(x) = 1/(1+x^2)$ *como una serie geométrica.*

b. *Intégrese la serie en* **a** *para demostrar*

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots, \quad |x| < 1.$$

c. *Justifíquese la fórmula de Euler:*

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots.$$

Solución

a. Desarrollamos $1/(1+x^2)$ como una serie geométrica:

$$\frac{1}{1+x^2} = \frac{1}{1-(-x^2)} = 1 + (-x^2) + (-x^2)^2 + (-x^2)^3 + \cdots$$
$$= 1 - x^2 + x^4 - x^6 + \cdots,$$

lo que es válido si $|-x^2| < 1$, es decir, si $|x| < 1$. La convergencia es uniforme en $[-1+\varepsilon, 1-\varepsilon]$ para todo $\varepsilon > 0$, por el criterio M.

b. Integramos de 0 a x para obtener

$$\int_0^x \frac{dt}{1+t^2} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots;$$

pero sabemos que la integral de $1/(1+t^2)$ es $\tan^{-1} t$, por lo que

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots \quad \text{para } |x| < 1.$$

c. Si hacemos $x = 1$ y usamos el hecho de que $\tan^{-1} 1 = \pi/4$, obtenemos la fórmula de Euler:

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots;$$

pero esto no está justificado del todo, pues la serie de $\tan^{-1} x$ es válida solamente para $|x| < 1$ (es, sin embargo, plausible, ya que $1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$ que es una serie alternante, converge). Para justificar la fórmula de Euler, podemos utilizar la forma finita del desarrollo de una serie geométrica:

$$\frac{1}{1+t^2} = 1 - t^2 + t^4 + \cdots + (-1)^n t^{2n} + (-1)^{n+1}\frac{t^{2n+2}}{1+t^2}.$$

Si integramos de 0 a 1, obtenemos

$$\frac{\pi}{4} = \tan^{-1} 1 = 1 - \frac{1}{3} + \frac{1}{5} - \cdots + \frac{(-1)^n}{2n+1} + (-1)^{n+1}\int_0^1 \frac{t^{2n+2}}{1+t^2}\,dt.$$

El proceso estará completo si podemos probar que el último término tiende a cero cuando $n \to \infty$. Pero

$$0 \le \int_0^1 \frac{t^{2n+2}}{1+t^2}\,dt \le \int_0^1 t^{2n+2}\,dt = \frac{1}{2n+3} \to 0 \text{ cuando } n \to \infty. \quad \blacklozenge$$

5.3.10 Ejemplo *Supongamos que $a_0, a_1, a_2, a_3, \ldots$ es una sucesión acotada de números reales. Muéstrese que la serie*

$$\sum_{k=0}^{\infty} \frac{a_k}{k!} x^k$$

converge en $\mathbb{R}$ a una función derivable $f(x)$ y que

$$f'(x) = \sum_{k=0}^{\infty} \frac{a_{k+1}}{k!} x^k.$$

Solución Ya vimos en el ejemplo **5.2.6** que la serie $\sum_{k=0}^{\infty}(a_k/k!)x^k$ converge en $\mathbb{R}$ a una función continua $f(x)$, considerando las funciones $g_k(x) = (a_k/k!)x^k$ en intervalos de longitud finita. Pero $g_k'(x) = (ka_k/k!)x^{k-1} = (a_k/(k-1)!)x^{k-1}$, luego

$$\sum_{k=0}^{\infty} g_k'(x) = \sum_{k=1}^{\infty} \frac{a_k}{(k-1)!} x^{k-1} = \sum_{k=0}^{\infty} \frac{a_{k+1}}{k!} x^k.$$

El argumento empleado en el ejemplo **5.2.6** se puede aplicar de nuevo para demostrar que esta serie converge uniformemente en cualquier intervalo de longitud finita a una función $g(x)$. Podemos aplicar el corolario **5.3.4** para demostrar que f es derivable y que $f'(x) = g(x)$. ♦

Ejercicios de §5.3

1. Analícese la validez del teorema **5.3.1** para la sucesión f_n dada por

$$f_n(x) = \frac{nx}{1 + nx^2}, \quad 0 \leq x \leq 1.$$

2. Muéstrese que la sucesión $\{f_n\}$ dada por $f_n(x) = n^3x^n(1 - x)$ converge puntualmente a $f = 0$ en $[0, 1]$ y utilícese entonces el teorema **5.3.1** para demostrar que la convergencia no es uniforme.

3. Analícese la validez de los teoremas **5.3.1** y **5.3.3** para $f_n(x) = \sqrt{n}x^n(1 - x)$ en $[0, 1]$. [Sugerencia: encuéntrese el máximo de $f_n(x)$.]

4. Verifíquese que $\int_0^x \operatorname{sen} t \, dt = 1 - \cos x$, usando

$$\operatorname{sen} x = \sum_0^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}.$$

5. Verifíquese que sen$'$ $x = \cos x$, empleando series.

6. Exprésese la suma de la serie $\sum_{n=1}^{\infty} x^n/n^2$ como una integral. ¿Qué ocurre si $x = -1$?

§5.4 Las funciones elementales

Algunas de las más importantes funciones del análisis matemático se conocen con frecuencia como "funciones elementales". Entre ellas están:

1. Las funciones exponenciales, como e^x, 2^x y 10^x.

2. Los logaritmos correspondientes.

3. Las funciones potenciales, como x^2, $x^{13/17}$ y, más misteriosamente, x^{π}, además de las funciones estrechamente relacionadas, como los polinomios y las funciones racionales.

4. Las funciones trigonométricas, como sen x y $\cos x$.

El cálculo de estas funciones se desarrolla, por lo general, en un primer curso de cálculo y no es necesario que lo repitamos aquí (de hecho, hemos supuesto que el lector ya sabe algo de ellas). Sin embargo, vamos a señalar algunos de los puntos donde pueden aparecer dificultades y a mostrar la forma en que las ideas de las últimas secciones se pueden aplicar en este contexto.

Funciones exponenciales y logarítmicas

Los estudiantes se encuentran con la exponenciación muy rápidamente en su experiencia matemática, cuando aprenden a abreviar el producto repetido de n copias de un número b como b^n. Se ve entonces que estos exponentes enteros satisfacen la propiedad fundamental

$$b^n \cdot b^k = b^{n+k}.$$

La extensión de esta propiedad lleva rápidamente a la útil definición de la exponenciación con exponentes enteros negativos y también con exponentes racionales:

$$b^{-n} = \frac{1}{b^n} \quad \text{y} \quad b^{p/q} = \sqrt[q]{b^p} = (\sqrt[q]{b})^p.$$

No hay problema en estos resultados, al menos cuando b es positivo, excepto por la posible duda de la existencia de raíces q-ésimas de números reales positivos b. Ya resolvimos ese problema en el capítulo 1, con el estudio de la completitud de $\mathbb{R}$; podemos definir la raíz q-ésima de b como el supremo del conjunto de números reales cuya potencia q-ésima es menor que b.

Surge otro problema sutil cuando consideramos los exponentes irracionales. ¿Cuál debería ser el significado de 2^π? Quisiéramos tener una función de x que se comportase bien, denotada 2^x, tal que $2^{p/q} = (\sqrt[q]{2})^p$ si x es un número racional p/q. Una forma de atacar este problema es observar que $\mathbb{Q}$ es denso en $\mathbb{R}$, de modo que cualquier número real, como π, se puede aproximar mediante números racionales. Podríamos seleccionar una sucesión $r_1, r_2, r_3, \ldots$ en $\mathbb{Q}$ tal que $r_n \to \pi$ cuando $n \to \infty$. Sabemos el significado que queremos para $2^{r_1}, 2^{r_2}, 2^{r_3}, \ldots$ y queremos que 2^x sea una función continua de x, por lo que deberíamos tener $2^\pi = \lim_{n\to\infty} 2^{r_n}$,. Ésta es una buena idea y una forma bastante razonable de proceder para obtener una aproximación numérica de 2^π. Presenta las siguientes dificultades como base para una definición:

1. ¿Por qué existe el límite?

2. ¿Por qué el límite es independiente de la elección de una sucesión de racionales que aproximan π?

3. Si hacemos lo mismo para x un real arbitrario, ¿cómo establecer la derivabilidad y demás aspectos para la función resultante 2^x?

4. ¿Cómo establecer la "ley de los exponentes" como $2^x 2^y = 2^{x+y}$?

Estos problemas no son insuperables, pero tampoco triviales.

La forma más común de atacarlos en un curso de cálculo comienza por el logaritmo, en vez de la función exponencial. El logaritmo natural se introduce por medio del teorema fundamental del cálculo como una primitiva de la función recíproca:

$$\ln(x) = \int_1^x \frac{1}{t}\,dt \quad \text{para } x > 0.$$

Esto hace automáticamente que el logaritmo sea una función derivable de x en la semirrecta $x > 0$, con derivada igual a $1/x$. Las propiedades fundamentales del logaritmo son una consecuencia elemental de esta definición. La función exponencial se introduce luego como la inversa del logaritmo. Aunque puede no ser muy intuitivo, este enfoque tiene varias ventajas. Los ejercicios del 2 al 7 del final de esta sección guían al lector a través de este desarrollo.

Aquí presentaremos un enfoque que comienza con la función exponencial, más conocida, pero definida en términos de una serie infinita. Es probable que el lector esté familiarizado con el desarrollo de las funciones elementales mediante una serie de potencias, conocida como la serie de Taylor o la serie de Maclaurin. En particular, un hecho ya conocido es que la función exponencial se puede expresar como

$$e^x = \sum_{k=0}^{\infty} \frac{1}{k!}x^k = 1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \cdots.$$

para todo x real.

Nuestro plan es usar esta serie para definir una función, que, según mostraremos, se comporta como lo haría la función exponencial.

5.4.1 Definición *Para cada número x, sea* $\exp(x) = \sum_{k=0}^{\infty} \frac{1}{k!}x^k$.

De **5.3.10**, esta serie converge uniforme y absolutamente en cualquier subconjunto acotado de $\mathbb{R}$ (o de C) y define una función derivable en todo $\mathbb{R}$. Podemos usar los resultados conocidos de las series para mostrar que se cumplen las leyes fundamentales de las exponenciales. Esto se sigue de la convergencia absoluta de la serie, pues podemos reordenar los términos del producto de dos series para agrupar los términos de potencias semejantes:

$$\begin{aligned}\left(1+x+\frac{x^2}{2}+\frac{x^3}{6}+\cdots\right)\left(1+y+\frac{y^2}{2}+\frac{y^3}{6}+\cdots\right)\\ &= 1+(x+y)+\frac{x^2+2xy+y^2}{2}+\frac{x^3+3x^2y+3xy^2+y^3}{6}+\cdots\\ &= 1+(x+y)+\frac{(x+y)^2}{2}+\frac{(x+y)^3}{6}+\cdots.\end{aligned}$$

Si escribimos esto con precisión mediante la notación del sumatorio y usando la fórmula del binomio para desarrollar $(x+y)^n$, obtenemos la "ley de los exponentes".

5.4.2 Proposición $\exp(x)\cdot\exp(y) = \exp(x+y)$ *para todo par de números x e y.*

Demostración
$$\begin{aligned}\exp(x)\exp(y) &= \left(\sum_{k=0}^{\infty}\frac{1}{k!}x^k\right)\left(\sum_{j=0}^{\infty}\frac{1}{j!}y^j\right)\\ &= \sum_{n=0}^{\infty}\left(\sum_{k+j=n}^{\infty}\frac{1}{k!}\frac{1}{j!}x^ky^j\right)\\ &= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}\frac{1}{k!}\frac{1}{(n-k)!}x^ky^{n-k}\right)\\ &= \sum_{n=0}^{\infty}\left(\frac{1}{n!}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}x^ky^{n-k}\right)\\ &= \sum_{n=0}^{\infty}\frac{1}{n!}(x+y)^n = \exp(x+y).\end{aligned}$$

Con este punto de partida, se siguen rápidamente muchas consecuencias.

5.4.3 Proposición exp(x) *es una función derivable en* $\mathbb{R}$ *tal que*

i. $\exp(0) = 1$.

ii. $\dfrac{d}{dx}(\exp(x)) = \exp(x)$.

iii. $\exp(x) > 0$ *para todo* x *en* $\mathbb{R}$.

iv. $\exp(-x) = 1/\exp(x)$.

v. $\exp(x)$ *es estrictamente creciente en todo* $\mathbb{R}$.

vi. $\exp(x) \to +\infty$ si $x \to +\infty$.

vii. $\exp(x) \to 0$ si $x \to -\infty$.

viii. exp *es una biyección de* $\mathbb{R}$ *sobre.* $\mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}$.

La derivabilidad está dada por la proposición **5.3.10.** La primera afirmación es obvia de la definición de exp(x) mediante una serie. La segunda es una consecuencia sencilla, pues la derivación término a término está permitida por **5.3.10**:

$$\begin{aligned}\frac{d}{dx}(\exp(x)) &= \frac{d}{dx}\left(1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \cdots\right)\\ &= 0 + 1 + x + \frac{3}{3!}x^2 + \frac{4}{4!}x^3 + \frac{5}{5!}x^4 + \cdots\\ &= 1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \cdots = \exp(x).\end{aligned}$$

De **5.4.2**, sabemos que $\exp(-x) \cdot \exp(-x) = \exp(-x + x) = \exp(0) = 1$. Así que $\exp(x) = 1/\exp(-x)$. En particular, exp(x) nunca se anula. Como $\exp(0) = 1$, debemos tener $\exp(x) > 0$ para todo x real. En caso contrario, la función continua exp tendría una raíz real, por el teorema de los valores intermedios. Como $\exp'(x) = \exp(-x) > 0$ en todo punto, concluimos que exp(x) es estrictamente creciente en $\mathbb{R}$. Esto implica que $\exp(t) > 1$ para todo $t > 0$. Mediante el teorema fundamental del cálculo, podemos obtener que, para $x > 0$,

$$\exp(x) = \exp(0) + \int_0^x \exp'(t)\, dt > 1 + \int_0^x 1\, dt = 1 + x.$$

Esto muestra que $\exp(x) \to +\infty$ si $x \to +\infty$. Como $\exp(-x) = 1/\exp(x)$, también tenemos que $\exp(x) \to 0$ si $x \to -\infty$. La combinación de **v**, **vi** y **vii** con el teorema de los valores intermedios muestra que exp aplica $\mathbb{R}$ de forma biyectiva sobre $\mathbb{R}^+$. La continuidad y el teorema de los valores intermedios se usan para obtener la suprayectividad.

Las proposiciones **5.4.2** y **5.4.3** indican que la función exp actúa como la exponenciación. Ahora mostraremos que esto es cierto. En primer lugar, definimos el número especial e:

$$e = \exp(1) = \sum_{k=0}^{\infty} \frac{1}{k!} = 1 + 1 + \frac{1}{2} + \frac{1}{3!} + \cdots.$$

Por medio de los logaritmos, que se presentan en el siguiente apartado y con la regla de l'Hôpital, se puede mostrar que $e = \text{lím}_{n\to\infty}(1 + (1/n))^n$. Esto proporciona un vínculo con temas como el interés compuesto "continuo". A partir de cualquiera de estas caracterizaciones o de otras, se puede calcular el valor de e con el grado de precisión deseado. Las primeras cifras decimales son

$$e \approx 2.71828\ 18284\ 59045\ 23536\ 02874\ 71352\ 66249\ \ldots$$

No debe sorprendernos que e sea irracional, pero incluso cumple otra condición. Ni siquiera es algebraico. Un número algebraico es un número que es raíz de un polinomio con coeficientes racionales. Los números que no son algebraicos son *trascendentes*. El hecho de que e es trascendente fue demostrado por Charles Hermite en 1873. (C.L.F. Lindemann llegó a la misma conclusión para π en 1882.)

La función $\exp(x)$ actúa igual que lo haría e elevado a la "potencia" x. En primer lugar, supóngase que n es un entero positivo. La aplicación de **5.4.2** varias veces da lugar a $\exp(n) = \exp(1 + 1 + \cdots + 1) = \exp(1) \cdot \exp(1) \cdots \exp(1) = (\exp(1))^n = e^n$ (el lector debe poder proporcionar la demostración de esto por inducción). Para los enteros negativos, utilizamos **5.4.3iv** para concluir que $\exp(-n) = 1/\exp(n) = 1/e^n$. Esto es exactamente lo que estábamos acostumbrados a escribir como e^{-n}. En forma análoga,

$$\begin{aligned}(\exp(1/n))^n &= \exp(1/n) \cdot \exp(1/n) \cdots \exp(1/n) \\ &= \exp((1/n) + (1/n) + \cdots + (1/n)) = \exp(1) = e.\end{aligned}$$

Así $\exp(1/n)$ es una raíz n-ésima de e. Estamos acostumbrados a escribir esto como $\exp(1/n) = \sqrt[n]{e} = e^{1/n}$. Juntamos todos estos hechos y vemos que para los enteros p y q, con $q \neq 0$, $\exp(p/q) = (\sqrt[q]{e})^p$, exactamente lo que estábamos acostumbrados a escribir como $e^{p/q}$.

5.4.4 Proposición $\exp(x)$ *es una función derivable en* $\mathbb{R}$ *tal que* $\exp(x) = e^x$ *para* $x \in \mathbb{Q}$.

La función $\exp(x)$ sirve entonces como una definición bastante razonable para el significado de e^x para x real (o, de hecho, para cualquier número complejo). Ahora tenemos una buena definición de e^π, aunque nuestro reto original era 2^π. Antes de enfrentar este problema, centraremos nuestra atención en los logaritmos.

Logaritmos naturales

La función exponencial $\exp(x) = e^x$ ha sido definida de modo que produzca una función derivable biyectiva de $\mathbb{R}$ en $\mathbb{R}^+$. Es estrictamente creciente, con una derivada positiva para todo número real x. En particular, la derivada nunca se anula. Por el teorema de la función inversa en una variable (**4.7.15**), existe una función inversa derivable que lleva $\mathbb{R}^+$ de forma biyectiva sobre $\mathbb{R}$, a la que llamamos función ***logaritmo natural*** y que denotamos $\ln x$ o $\log x$. El hecho de que sea la inversa de la exponencial quiere decir que

$$\log(e^x) = x \quad \text{para todo } x \text{ real}$$

y

$$\exp(\log x) = x \quad \text{para todo } x > 0.$$

El cálculo con la regla de la cadena nos produce la derivada:

$$\frac{d}{dx}(\exp(\log x)) = \frac{d}{dx}(x), \qquad \text{es decir,} \qquad \exp(\log x)\frac{d}{dx}(\log x) = 1.$$

Así,

$$x\frac{d}{dx}(\log x) = 1, \qquad \text{es decir,} \qquad \frac{d}{dx}(\log x) = \frac{1}{x}.$$

Así, el logaritmo natural es una primitiva de la función recíproca. Como $\log(1) = 0$, el teorema fundamental del cálculo implica que

$$\log x = \int_1^x \frac{1}{t}\,dt \quad \text{para } x > 0,$$

la fórmula que comúnmente se toma como punto de partida para la teoría en un curso de cálculo, como ya hemos mencionado.

Otras bases

Una vez con el logaritmo natural a nuestro alcance, podemos dedicar nuestra atención al reto original. ¿Cuánto vale 2^π? Si nos adelantamos un poco y suponemos que esto y el logaritmo natural se comportan como deben, podemos realizar un cálculo que nos indicará cuál debe ser la forma correcta de la definición. Quisiéramos que $\log(2^\pi) = \pi \log 2$, de modo que $2^\pi = \exp(\log(2^\pi)) = \exp(\pi \log 2)$. Como exp y log son funciones conocidas, esta última expresión proporciona la base para una definición. Podemos intentar definir 2^π como $\exp(\pi \log 2)$, o, más en general, 2^x como $\exp(x \log 2)$ para cualquier x. Mostraremos que esto funciona no sólo para 2, sino para cualquier número positivo b como base para dar una definición razonable de b^x.

5.4.5 Definición *Sea $b > 0$. Para cualquier número x, se define*

$$\exp_b(x) = \exp(x \log b).$$

Aplicamos las propiedades ya conocidas de exp y la regla de la cadena para obtener las siguientes propiedades básicas de $\exp_b$:

5.4.6 Proposición *Si $b > 0$, entonces $\exp_b(x)$ es una función derivable en $\mathbb{R}$ y también:*

i. $\exp_b(x + y) = \exp_b(x) \cdot \exp_b(y)$ *para todo par de números x e y.*

ii. $\exp_b(0) = 1$.

iii. $\dfrac{d}{dx}(\exp_b(x)) = \log(b) \cdot \exp_b(x)$ *para todo número real x.*

iv. $\exp_b(x) > 0$ *para todo número real x.*

v. $\exp_b(-x) = 1/\exp_b(x)$.

Si $b = 1$, entonces $\log(b) = 0$. Para $0 < b < 1$, tenemos que $\log b < 0$, mientras que si $b > 1$, entonces $\log b > 0$. Como $\exp_b(t) > 0$ para todo t y $\exp_b'(x) = \log(b) \cdot \exp_b(x \log b)$, esto nos da toda la información necesaria sobre el signo de $\exp_b'(x)$.

5.4.7 Proposición

i. *Si $b = 1$, entonces $\exp_b(x) = 1$ para todo $x \in \mathbb{R}$.*

ii. *Si $b > 1$, entonces*

- **a.** $\exp_b(x)$ *es estrictamente decreciente para todo $x \in \mathbb{R}$.*
- **b.** $\exp_b(x) \to +\infty$ *si* $x \to +\infty$.
- **c.** $\exp_b(x) \to 0$ *si* $x \to -\infty$.

iii. *Si $0 < b < 1$, entonces*

- **a.** $\exp_b(x)$ *es estrictamente creciente para todo $x \in \mathbb{R}$.*
- **b.** $\exp_b(x) \to 0$ *si* $x \to +\infty$.
- **c.** $\exp_b(x) \to +\infty$ *si* $x \to -\infty$.

En los casos **ii** *y* **iii**, *la función $\exp_b(x)$ es una biyección de $\mathbb{R}$ sobre $\mathbb{R}^+$.*

El cálculo que muestra la coincidencia de $\exp_b(x)$ con nuestra noción usual de b^x para x racional es análogo al que hicimos para $\exp(x)$.

5.4.8 Proposición *$\exp_b(x)$ es una función continua y derivable en $\mathbb{R}$ tal que $\exp_b(x) = b^x$ para todo $x \in \mathbb{Q}$.*

Así, $\exp_b(x)$ sirve como una definición razonable de b^x para cualquier número x. Por supuesto, como $\exp_b(x)$ es una biyección de $\mathbb{R}$ sobre $\mathbb{R}^+$ (al menos para $b \neq 1$), cuya derivada nunca se anula, existe una función inversa derivable,

$$\log_b(b^x) = x \quad \text{para todo } x \in \mathbb{R},$$

que es una biyección de $\mathbb{R}^+$ sobre $\mathbb{R}$, y

$$\exp_b(\log_b(x)) = x \quad \text{para todo } x > 0.$$

Ahora podemos escribir las propiedades básicas de las exponenciales y los logaritmos. En primer lugar, usamos la base e para calcular $\log(b^x) = \log(\exp(x \log b)) = x \log b$. Así, $(b^x)^y = \exp(y \log(b^x)) = \exp(yx \log b) = \exp(xy \log b) = b^{xy}$. Ésta es la última de las principales propiedades algebraicas de la exponenciación.

5.4.9 Proposición *Si $b > 0$ y x e y están en $\mathbb{R}$, entonces*

i. $b^0 = 1$.

ii. $b^x > 0$ *para todo x en $\mathbb{R}$.*

iii. $b^x b^y = b^{x+y}$.

iv. $(b^x)^y = b^{xy}$.

Ahora disponemos también de las propiedades correspondientes de los logaritmos:

5.4.10 Proposición *Si $b > 0$ y $b \neq 1$, entonces*

i. $\log_b(1) = 0$.

ii. $\log_b(xy) = \log_b(x) + \log_b(y)$ *para todo x e y positivos.*

iii. $\log_b(x^t) = t \log_b(x)$ *para $x > 0$ y $t \in \mathbb{R}$.*

Para obtener **ii**, sean $x = b^s$ e $y = b^t$ y calculemos

$$\log_b(xy) = \log_b(b^s b^t) = \log_b(b^{s+t}) = s + t = \log_b(b^s) + \log_b(b^t)$$
$$= \log_b(x) + \log_b(y).$$

Para obtener **iii**, sea $x = b^s$ y calculemos

$$\log_b(x^t) = \log_b(b^s)^t = \log_b(b^{st}) = st = t\log_b(b^s) = t\log_b(x).$$

El cálculo de estas funciones se resume en la siguiente proposición.

5.4.11 Proposición

i. $\dfrac{d}{dx}(e^x) = e^x; \int e^x\,dx = e^x + C.$

ii. $\dfrac{d}{dx}(\log x) = 1/x; \int_1^x (1/t)\,dt = \log x, \ x > 0.$

iii. $\dfrac{d}{dx}(b^x) = b^x \log b; \int b^x\,dx = b^x/\log b + C.$

iv. $\dfrac{d}{dx}(\log_b x) = 1/(x\log b).$

Funciones potenciales

Por una "función potencial" entendemos una función de la forma $f(x) = x^p$ para una constante p. Las únicas funciones potenciales que realmente hemos usado hasta ahora en el desarrollo de la teoría son las que tienen p entero, aunque hemos analizado en cierta medida las raíces, correspondientes a valores racionales de p. Ahora tenemos una buena definición de x^p para cualquier p, al menos si x es positivo:

$$x^p = \exp(p\log x).$$

Sabemos que esto satisface la mayoría de las propiedades correctas de las potencias de números positivos:

$$x^p x^r = x^{p+r}$$
$$(x^p)^r = x^{pr}$$
$$x^{-p} = 1/x^p$$
$$x^0 = 1 \quad \text{si } x \neq 0.$$

Una propiedad adicional, $x^p y^p = (xy)^p$, se puede verificar fácilmente:

$$(xy)^p = \exp(p\log xy) = \exp(p(\log x + \log y)) = \exp(p\log x + p\log y)$$
$$= \exp(p\log x)\cdot\exp(p\log y) = x^p y^p.$$

La regla familiar para las derivadas de las funciones potenciales se obtiene rápidamente de la regla de la cadena:

$$\frac{d}{dx}(x^p) = \frac{d}{dx}(\exp(p\log x)) = \exp'(p\log x)\cdot\frac{d}{dx}(p\log x)$$
$$= \exp(p\log x)\cdot p/x = (x^p)\cdot p/x = px^{p-1}.$$

Por supuesto, esto funciona en toda su generalidad si $x > 0$. Sin embargo, siempre podemos hacer $0^p = 0$ para cada $p > 0$ y

$$x^{p/q} = \sqrt[q]{x^p}$$

para todo número real x si p y q son enteros y p es par.

Funciones trigonométricas

Es probable que el lector conozca los desarrollos de las funciones trigonométricas seno y coseno:

$$\operatorname{sen} x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

y

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots.$$

Cada una de estas series es de la forma $\sum_{k=0}^{\infty}(a_k/k!)x^k$, donde cada uno de los coeficientes a_k es -1, 0 o 1. Así, se ajustan al patrón del ejemplo **5.3.10**. Así, si definimos las funciones $s(x)$ y $c(x)$ mediante las series

$$s(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \cdots = \sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}x^{2k+1}$$

y

$$c(x) = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \cdots = \sum_{k=0}^{\infty}\frac{(-1)^k}{2k!}x^{2k},$$

cada una de éstas converge uniforme y absolutamente en cualquier conjunto acotado de números y representa una función derivable en todo $\mathbb{R}$. La derivación término a término

es válida para todo número real x. Si derivamos, obtenemos

$$s'(x) = c(x) \quad \text{y} \quad c'(x) = -s(x) \quad \text{para todo número real } x.$$

De esta serie y de las relaciones entre las derivadas, podemos mostrar que $s(x)$ y $c(x)$ satisfacen las mismas identidades fundamentales que el seno y el coseno. Por ejemplo, sea $h(x) = s(x)^2 + c(x)^2$. Entonces $h(x)$ es derivable y $h'(x) = 2s(x)c(x) - 2c(x)s(x) = 0$ para todo x. Así, $h(x)$ es constante. Como $c(0) = 1$ y $s(0) = 0$, concluimos que $h(x) = 1$ para todo x. Es decir,

$$s(x)^2 + c(x)^2 = 1 \quad \text{para todo número real } x.$$

También podemos ver de la serie que $s(x)$ es una función impar de x y que $c(x)$ es par:

$$s(-x) = -s(x) \quad \text{y} \quad c(-x) = c(x) \quad \text{para todo } x.$$

Las fórmulas para el seno y el coseno de una suma también se pueden recuperar mediante la manipulación directa de la serie, casi de la misma forma en que obtuvimos la fórmula fundamental para la función exponencial, $\exp(x + y) = \exp(x) \cdot \exp(y)$. Sin embargo, también podemos obtenerlas de forma indirecta, observando un vínculo importante entre la función exponencial y las funciones trigonométricas. Esta relación depende de la convergencia uniforme y absoluta de la serie $\sum_{k=0}^{\infty}(a_k/k!)x^k$ en cualquier conjunto acotado de números *complejos* siempre que los coeficientes *complejos* estén acotados. La demostración sigue el mismo método que para los números reales, al igual que las manipulaciones que nos condujeron a la fórmula $\exp(x + y) = \exp(x) \cdot \exp(y)$. Si calculamos $\exp(ix)$ para x real, tenemos que

$$\begin{aligned} e^{ix} &= \sum_{k=0}^{\infty} \frac{1}{k!}(ix)^k \\ &= 1 + ix - \frac{x^2}{2} - \frac{ix^3}{3} + \frac{x^4}{4} + \frac{ix^5}{5} - \frac{x^6}{6} - \cdots \\ &= \left(1 - \frac{x^2}{2} + \frac{x^4}{4} - \cdots\right) + i\left(x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots\right) \\ &= c(x) + is(x). \end{aligned}$$

La ley de los exponentes nos garantiza que $e^{i(x+y)} = e^{ix}e^{iy}$, de modo que podemos calcular

$$\begin{aligned} c(x+y) + is(x+y) &= [c(x) + is(x)] \cdot [c(y) + is(y)] \\ &= [c(x)c(y) - s(x)s(y)] + i[s(x)c(y) + c(x)s(y)]. \end{aligned}$$

Si comparamos las partes real e imaginaria, obtenemos

$$c(x+y) = c(x)c(y) - s(x)s(y) \quad \text{y} \quad s(x+y) = s(x)c(y) + c(x)s(y).$$

Las funciones definidas por la serie son entonces funciones derivables en todo $\mathbb{R}$ que satisfacen las mismas identidades fundamentales que las funciones trigonométricas. Parece casi un hecho probado que deben ser las mismas funciones. De hecho, lo son, pero el vínculo no se establece tan directamente como en el caso de la exponencial. Nuestro conocimiento previo de la exponencial era esencialmente algebraico: $e^{p/q} = \sqrt[q]{e^p}$. La confirmación de que esto era lo mismo que $\exp(p/q)$ era una rápida consecuencia del manejo algebraico de la serie. Sin embargo, las definiciones usuales del seno y el coseno se basan firmemente en la geometría de los ángulos o en la geometría de la longitud de arco a lo largo de una circunferencia. Si θ es la longitud de un arco a lo largo de la circunferencia de radio 1 con centro en $(0, 0)$ desde $(1, 0)$ hasta un punto (α, β) sobre la circunferencia, entonces $\alpha = \cos(\theta)$ y $\beta = \operatorname{sen}(\theta)$. Véase la figura 5.4.1.

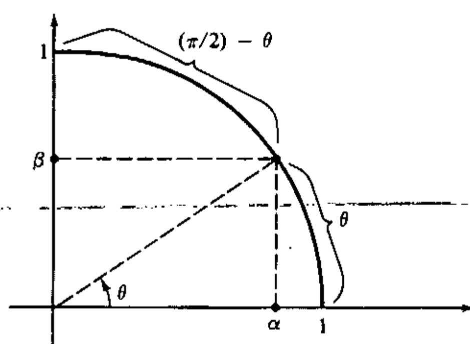


FIGURA 5.4-1 El seno y coseno de θ son α y β

En un primer curso de cálculo, lo común es que los límites básicos

$$\lim_{\theta\to 0} \frac{\operatorname{sen}\theta}{\theta} = 1 \qquad \text{y} \qquad \lim_{\theta\to 0} \frac{\cos\theta - 1}{\theta} = 0$$

se demuestren geométricamente. Las fórmulas para el seno y el coseno de la suma de dos ángulos se presuponen y se utilizan con estos dos límites para mostrar que seno y coseno son derivables y que

$$\frac{d}{d\theta}(\operatorname{sen}\theta) = \cos\theta \qquad \text{y} \qquad \frac{d}{d\theta}(\cos\theta) = -\operatorname{sen}\theta.$$

Este proceso se puede invertir si usamos la fórmula para la longitud de la gráfica de una función derivable,

$$\text{Longitud} = \int_a^b \sqrt{1 + [f'(x)]^2}\, dx.$$

El primer cuadrante de la circunferencia $x^2 + y^2 = 1$ se puede escribir como la gráfica de las siguientes funciones en ambas direcciones:

$$y = f(x) = \sqrt{1 - x^2} \qquad \text{y} \qquad x = g(y) = \sqrt{1 - y^2}.$$

Esto nos lleva a

$$\frac{\pi}{2} - \theta = \int_0^{\alpha} \sqrt{1 + [f'(x)]^2}\, dx \qquad \text{y} \qquad \theta = \int_0^{\beta} \sqrt{1 + [g'(x)]^2}\, dy$$

o

$$\theta = \frac{\pi}{2} - \int_0^{\alpha} \frac{1}{\sqrt{1 - x^2}}\, dx \qquad \text{y} \qquad \theta = \int_0^{\beta} \frac{1}{\sqrt{1 - y^2}}\, dy.$$

Las funciones inversas $\theta = \arccos \alpha$ y $\theta = \operatorname{arcsen} \beta$ son entonces funciones derivables de α y β y el teorema fundamental del cálculo implica que

$$\frac{d\theta}{d\alpha} = -\frac{1}{\sqrt{1 - \alpha^2}} \qquad \text{y} \qquad \frac{d\theta}{d\beta} = \frac{1}{\sqrt{1 - \beta^2}}.$$

El teorema de la función inversa dice que α y β son funciones derivables de θ con

$$\frac{d\alpha}{d\theta} = -\sqrt{1 - \alpha^2} \qquad \text{y} \qquad \frac{d\beta}{d\theta} = \sqrt{1 - \beta^2}.$$

Recordando que $\alpha^2 + \beta^2 = 1$, tenemos

$$\frac{d\alpha}{d\theta} = -\beta \qquad \text{y} \qquad \frac{d\beta}{d\theta} = \alpha$$

o

$$\frac{d}{d\theta} \cos \theta = -\operatorname{sen} \theta \qquad \text{y} \qquad \frac{d}{d\theta} \operatorname{sen} \theta = \cos \theta.$$

Los pares de funciones

$$\begin{array}{ll} x(t) = c(t) & x(t) = \cos t \\ y(t) = s(t) & y(t) = \operatorname{sen} t \end{array}$$

satisfacen ambos el sistema de ecuaciones de primer orden

$$x'(t) = -y(t) \qquad y'(t) = x(t)$$

con el mismo conjunto de condiciones iniciales

$$x(0) = 1 \qquad \text{e} \qquad y(0) = 0.$$

Un resultado básico de la teoría de ecuaciones diferenciales que demostraremos en §**7.5** muestra que la solución de un tal sistema es única. Concluimos que $c(t)$ es lo mismo que $\cos t$ y que $s(t)$ es lo mismo que $\operatorname{sen} t$.

No hemos obtenido la misma ganancia del estudio mediante series de las funciones trigonométricas que del de la exponencial. Sin embargo, permitiendo que el argumento tenga valores complejos, hemos descubierto un vínculo interesante entre ambas: $e^{ix} = \cos x + i \operatorname{sen} x$, y en la otra dirección,

$$\cos x = \frac{e^{ix} + e^{-ix}}{2} \quad \text{y} \quad \operatorname{sen} x = \frac{e^{ix} - e^{-ix}}{2i}.$$

Si $x = \pi$, encontramos una admirable ecuación que relaciona cinco de los más importantes números en matemáticas: $e^{i\pi} + 1 = 0$.

5.4.12 Ejemplo *Verifíquese que* $\int_0^x \cos\theta \, d\theta = \operatorname{sen} x$ *mediante series.*

Solución Sabemos que la serie

$$\cos\theta = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}\theta^{2k}$$

converge uniformemente en cualquier subconjunto acotado de $\mathbb{R}$. Si elegimos $b > |x|$, entonces tenemos la convergencia uniforme en $[-b, b]$, de modo que podemos integrar término a término entre 0 y x:

$$\int_0^x \cos\theta \, d\theta = \int_0^x \left(\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}\theta^{2k} \right) d\theta = \sum_{k=0}^{\infty} \left(\int_0^x \frac{(-1)^k}{(2k)!}\theta^{2k} \, d\theta \right)$$
$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}x^{2k+1} = \operatorname{sen} x. \quad \blacklozenge$$

Ejercicios de §5.4

1. Verifíquese que $\int_0^x \operatorname{sen} t \, dt = 1 - \cos x$, mediante su desarrollo en serie.

En los ejercicios del 2 al 7 pedimos al lector que desarrolle algunos aspectos de las exponenciales y los logaritmos a partir de la definición del logaritmo natural como primitiva de la función recíproca.

2. Sea $f(x) = \int_1^x (1/t) \, dt$ para $x > 0$. Muéstrese que f es una función derivable en $\mathbb{R}^+ = \{x \in \mathbb{R} \mid x > 0\}$ tal que $f(1) = 0$ y $f'(x) = 1/x$ para todo $x > 0$.

3. Muéstrese que si $a > 0$ y $b > 0$, entonces $f(ab) = f(a) + f(b)$ (f está definida en el ejercicio 2).

4. Muéstrese que

a. f es estrictamente creciente en $\mathbb{R}^+$ (f está definida en el ejercicio 2).

b. $\lim_{x\to\infty} f(x) = +\infty$.

c. $\lim_{x\to 0} f(x) = -\infty$.

5. Muéstrese que f tiene una función inversa derivable $g = f^{-1}$ que es una biyección de $\mathbb{R}$ sobre $\mathbb{R}^+$ y que $g'(x) = g(x)$ para todo x real (f está definida en el ejercicio 2).

6. Muéstrese que $g(x + y) = g(x) \cdot g(y)$ para todo real x e y (g está definida en el ejercicio 5).

7. Sea $e = g(1)$ y muéstrese que $g(p/q) = e^{p/q}$ para todo número racional p/q, con p y q en $\mathbb{Z}$ y $q \neq 0$.

8. La "función error" se define como

$$\operatorname{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2}\, dt.$$

a. Muéstrese que $\operatorname{erf}(x)$ se puede representar mediante una serie de potencias

$$\sum_{k=0}^{\infty} a_k x^k$$

válida para todo x, y calcule a_0, a_1, a_2, a_3, a_4, y a_5.

b. Utilícese el resultado principal del apartado **a** para estimar el valor de

$$\frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-x^2/2}\, dx.$$

(Esta integral da la probabilidad de que una medida tomada al azar de una población con distribución normal se separe como mucho una desviación estándar de la media 0.)

§5.5 El espacio de las funciones continuas

Fíjese un espacio métrico M, un subconjunto $A \subset M$ y un espacio vectorial normado N. Considérese el conjunto V de todas las funciones $f\colon A \to N$. Entonces, es fácil ver que V es un espacio vectorial. En V, el vector cero es la función que vale 0 para todo $x \in A$

y la suma y el producto por un escalar se definen como $(f+g)(x) = f(x) + g(x)$ y $(\lambda f)(x) = \lambda(f(x))$ para cada $\lambda \in \mathbb{R}$ y $f, g \in V$. Sea

$$C = \{f \in V \mid f \text{ es continua}\}$$

Si existe peligro de confusión, escribimos $C(A, N)$ para indicar que C depende de nuestra elección de A y N. Entonces C es también un espacio vectorial, pues la suma de dos funciones continuas es continua y un múltiplo escalar de una función continua es continuo.

Sea C_b el subespacio vectorial de C formado por las *funciones acotadas*: $C_b = \{f \in C \mid f \text{ está acotada}\}$. Recuérdese que "$f$ está acotada" significa que existe una constante M tal que $\|f(x)\| \leq M$ para todo $x \in A$. Si A es compacto, entonces $C_b = C$, por el teorema **4.4.1** aplicado a la función escalar $x \mapsto \|f(x)\|$.

Para $f \in C_b$, sea $\|f\| = \sup\{\|f(x)\| \mid x \in A\}$, el cual existe, pues f está acotada. El número $\|f\|$ es una medida del tamaño de f y se denomina ***norma*** de f. Véase la figura 5.5-1. Obsérvese que $\|f\| \leq M$ sup sii $\|f(x)\| \leq M$ para todo $x \in A$.

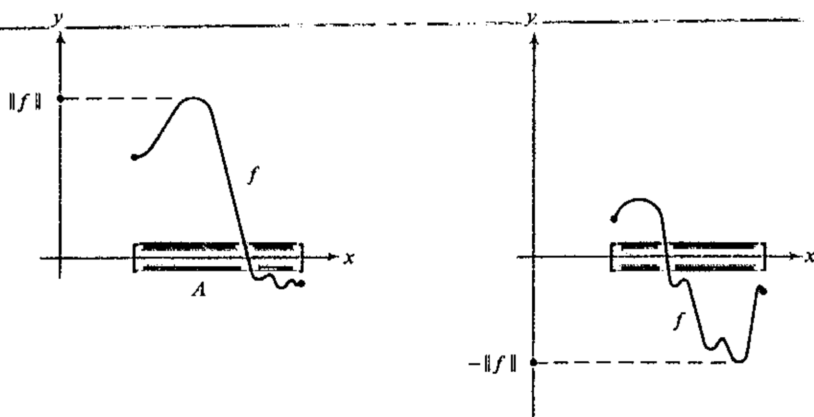


FIGURA 5.5-1 La norma de una función selecciona el máximo valor absoluto de $f(x)$

Lo que intentamos hacer aquí es analizar el espacio C_b de la misma forma en que analizamos $\mathbb{R}^n$. Es decir, cada punto de C_b (que es una función) tiene una norma, por lo que podemos esperar que muchos de los conceptos desarrollados para los vectores de $\mathbb{R}^n$ sean válidos en C_b. Tal punto de vista es útil en análisis, y se pueden demostrar algunos resultados importantes aplicando algunas de nuestras técnicas al espacio C_b. Para que este programa tenga éxito, la primera tarea es establecer que C_b es un *espacio normado*.

Nota. Aunque tenemos una norma, *no* tenemos un producto interno asociado con ella tal que $\|f\|^2 = \langle f, f\rangle$. Otros espacios de funciones, que estudiamos en el análisis de Fourier (capítulo 10) *sí* tienen un producto interno.

Si N es solamente un espacio métrico, aún seguimos teniendo una *métrica* en C_b, a saber,

$$d(f,g) = \sup\{\rho(f(x), g(x)) \mid x \in A\}.$$

La mayor parte de lo que sigue es válido en este contexto. Cuando sea necesaria una estructura de espacio vectorial en N lo estableceremos de forma explícita.

5.5.1 Teorema

i. *Si (M, d) y (N, ρ) son espacios métricos, entonces también lo es $C_b(A, N)$; es decir, $d(f, g)$ satisface*

a. $d(f,g) \geq 0$ *y* $d(f,g) = 0$ *sii* $f = g$.

b. $d(f,g) = d(g,f)$.

c. $d(f,g) \leq d(f,h) + d(h,g)$.

ii. *Si N es un espacio normado, entonces también lo es $C_b(A, N)$; es decir, $\|\cdot\|$ satisface*

a. $\|f\| \geq 0$ *y* $\|f\| = 0$ *sii* $f = 0$.

b. $\|\alpha f\| = |\alpha| \, \|f\|$ *para* $\alpha \in \mathbb{R}, f \in C_b$.

c. $\|f + g\| \leq \|f\| + \|g\|$ (***desigualdad triangular***).

Éstas son las reglas básicas necesarias para hablar de conjuntos abiertos, convergencia, etcétera. Por ejemplo, escribimos $f_k \to f$ en C_b sii $\|f_k - f\| \to 0$. La conexión con la convergencia uniforme es sencilla.

5.5.2 Teorema $(f_k \to f$ *(uniformemente en A)*$) \Leftrightarrow (f_k \to f$ *en* $C_b)$.

Recuérdese que una sucesión f_k es una ***sucesión de Cauchy*** si para cada $\varepsilon > 0$ existe N tal que $k, l \geq N$ implica $\|f_k - f_l\| < \varepsilon$. Recuérdese que un espacio es ***completo*** si toda sucesión de Cauchy converge. Otro nombre para un espacio normado completo es el de ***espacio de Banach***. La completitud es una propiedad técnica importante para un espacio, ya que con frecuencia podremos demostrar que una sucesión satisface el criterio de Cauchy y de ahí querríamos deducir que converge a algún elemento del espacio.

5.5.3 Teorema *Si N es un espacio métrico completo, entonces $C_b(A, N)$ es un espacio métrico completo. Si N es un espacio de Banach, también lo es $C_b(A, N)$.*

Éste resultado no es más que una transcripción de dos resultados básicos:

1. Una sucesión de funciones uniformemente de Cauchy en un espacio completo converge uniformemente a algo.

2. Un límite uniforme de funciones continuas es continuo.

El espacio C_b es sólo uno de muchos espacios de funciones de gran importancia para el análisis. Aunque C_b y $\mathbb{R}^n$ son espacios normados completos, son un poco diferentes en otros aspectos. Por ejemplo, como ya hemos mencionado, C_b no tiene un producto interno que produzca la norma $\|\cdot\|$ (véanse los ejercicios 12 y 30 al final del capítulo 1). Otra diferencia es que C_b no tiene dimensión finita. En las siguientes secciones veremos algunos problemas específicos a los que se puede aplicar esta teoría.

5.5.4 Ejemplo *Sea $B = \{f \in C([0, 1], \mathbb{R}) \mid f(x) > 0$ para todo $x \in [0, 1]\}$. Muéstrese que B es un conjunto abierto en $C([0, 1], \mathbb{R})$.*

Solución Por definición, para $f \in B$ debemos encontrar $\varepsilon > 0$ tal que $D(f, \varepsilon) = \{g \in C \mid \|f - g\| < \varepsilon\} \subset B$. Como $[0, 1]$ es compacto, f tiene un valor mínimo m en algún punto de $[0, 1]$. Así, $f(x) \geq m > 0$ para todo $x \in [0, 1]$. Sea $\varepsilon = m/2$. Si $\|f - g\| < \varepsilon$, entonces, para cada x, $\|f(x) - g(x)\| < \varepsilon = m/2$. Por lo tanto, $g(x) \geq m/2 > 0$ y entonces $g \in B$. ♦

5.5.5 Ejemplo *¿Cuál es la clausura del conjunto B del ejemplo* **5.5.4**?

Solución Afirmamos que la clausura es $D = \{f \in C \mid f(x) \geq 0$ para todo $x \in [0, 1]\}$. Éste es un conjunto cerrado, pues si $f_n(x) \geq 0$ y $f_n \to f$ uniformemente (y por lo tanto puntualmente), entonces $f(x) \geq 0$ para todo x. Para mostrar que D es la clausura, basta mostrar que para $f \in D$ existen $f_n \in B$ tales que $f_n \to f$ (¿por qué?). Para ello, defínase $f_n = f + 1/n$. ♦

5.5.6 Ejemplo *Considérese una sucesión $f_n \in C_b$ tal que $\|f_{n+1} - f_n\| \leq r_n$, donde $\sum r_n$ es convergente y $r_n \geq 0$. Demuéstrese que f_n converge.*

Solución Por la desigualdad triangular,

$$\begin{aligned}\|f_n - f_{n+k}\| &\leq \|f_n - f_{n+1}\| + \|f_{n+1} - f_{n+2}\| + \cdots + \|f_{n+k-1} - f_{n+k}\| \\ &\leq r_n + r_{n+1} + \cdots + r_{n+k}.\end{aligned}$$

Como $\sum r_i$ es convergente, esta expresión tiende a 0 cuando $n \to \infty$, puesto que es menor o igual que $s - s_{n-1}$, donde s_n es la n-ésima suma parcial y s es la suma. Por lo tanto, f_n es una sucesión de Cauchy, de modo que converge. ♦

Ejercicios de §5.5

1. Sea $B = \{f \in C_b(\mathbb{R}, \mathbb{R}) \mid f(x) > 0 \text{ para todo } x \in \mathbb{R}\}$. ¿Es B abierto? Si no lo es, ¿qué es $\text{int}(B)$?

2. ¿Cuál es la clausura de B en el ejercicio 1?

3. ¿Hay alguna relación entre el ejemplo **5.5.6** y el criterio M de Weierstrass? Razónese la respuesta.

4. Sea

$$f_n(x) = \frac{1}{n}\frac{nx}{1+nx}, \quad 0 \leq x \leq 1.$$

Muéstrese que $f_n \to 0$ en $C([0, 1], \mathbb{R})$.

5. Sea f_k una sucesión convergente en $C_b(A, \mathbb{R}^m)$. Demuéstrese que $\{f_k \mid k = 1, 2, \ldots\}$ está acotado en $C_b(A, \mathbb{R}^m)$. ¿Es cerrado?

§5.6 Teorema de Arzela-Ascoli

Este teorema es el análogo al teorema de Heine-Borel relativo a los subconjuntos compactos de $\mathbb{R}^n$ y da condiciones para que un conjunto en C_b sea compacto. Recuérdese que en $\mathbb{R}^n$ un conjunto es compacto si y sólo si es cerrado y acotado, pero que este criterio no necesariamente es válido en otros espacios métricos. En particular, no es válido en C_b. Lo que hace el teorema de Arzela-Ascoli es acercarse lo más posible a este resultado, con la intención de proporcionar condiciones para la compacidad que puedan verificarse en los ejemplos. Comenzaremos con algo de terminología.

5.6.1 Definición *Sea $\mathcal{B} \subset C(A, N)$. Decimos que $\mathcal{B}$ es un conjunto **equicontinuo** de funciones si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $x, y \in A$ y $d(x, y) < \delta$ implican $d(f(x), f(f)) < \varepsilon$ para todo $f \in \mathcal{B}$.*

Esta definición es igual a la de convergencia uniforme, excepto que ahora pedimos también que δ sea independiente de f, así como de x_0.

Sea $\mathcal{B}_x = \{ f(x) \mid f \in \mathcal{B}\}$ para x fijo. Éste es el conjunto de los valores de todas las funciones de $\mathcal{B}$ en el punto $x \in A$. Decimos que $\mathcal{B}$ es ***puntualmente compacto*** si y sólo si $\mathcal{B}_x$ es compacto en N para cada $x \in A$.

5.6.2 Teorema de Arzela-Ascoli *Sean $A \subset M$ compacto y $\mathcal{B} \subset C(A, N)$. Entonces $\mathcal{B}$ es compacto si y sólo si $\mathcal{B}$ es cerrado, equicontinuo y puntualmente compacto.*

La estrategia de la demostración se basa en la propiedad de Bolzano-Weierstrass. El punto principal es mostrar que si cada f_n está en $\mathcal{B}$, entonces la sucesión f_n tiene una subsucesión uniformemente convergente. Como $\mathcal{B}$ es puntualmente compacto, podemos hacer esto en *cada* $x \in A$. La idea entonces es usar la equicontinuidad y la compacidad de A para que esto sea uniforme. Un inteligente *proceso de selección diagonal* hace esto factible.

El teorema se utiliza con frecuencia, por ejemplo, como sigue:

5.6.3 Corolario *Sea $A \subset M$ compacto y $N = \mathbb{R}^m$. Supóngase que $\mathcal{B} \subset C(A, \mathbb{R}^m)$ es equicontinuo y puntualmente acotado. Entonces toda sucesión en $\mathcal{B}$ tiene una subsucesión uniformemente convergente.*

Aquí se usa el hecho de que cualquier conjunto acotado en $\mathbb{R}^m$ está dentro de un conjunto compacto (una bola, por ejemplo). Un resultado relacionado es el siguiente: Si $A \subset M$ es compacto y $N = \mathbb{R}^n$, entonces $\mathcal{B} \subset C(A, \mathbb{R}^m)$ *es compacto sii es acotado, cerrado y equicontinuo.* Dejaremos la demostración de esto al lector.

5.6.4 Ejemplo *Sean $f_n : [0, 1] \to \mathbb{R}$ continuas y tales que $|f_n(x)| \leq 100$ para todo n y todo x en $[0, 1]$; además, las derivadas f_n' existen y están uniformemente acotadas en $]0, 1[$. Demuéstrese que f_n tiene una subsucesión convergente.*

Solución Verificamos que el conjunto $\{ f_n\}$ es equicontinuo y acotado. La hipótesis es que $f_n'(x) \mid \leq M$ para una constante M. Por el teorema del valor medio,

$$|f_n(x) - f_n(y)| \leq M|x - y|,$$

y entonces, dado ε, podemos elegir $\delta = \varepsilon/M$ independiente de x, y y n. Así, $\{f_n\}$ es equicontinuo. Está acotado, pues $\|f_n\| = \sup_{0<x<1} |f_n(x)| \leq 100$. ♦

5.6.5 Ejemplo *¿Es cierto el resultado del ejemplo* **5.6.4** *si omitimos la condición de que* $|f_n(x)|$ *esté acotada?*

Solución No, ya que si $f_n(x) = n$, entonces $f_n' = 0$, pero es claro que no existe una subsucesión convergente. ♦

Para explotar la compacidad y combinar el teorema de Arzela-Ascoli con los resultados de funciones continuas del capítulo 4, necesitamos disponer de funciones continuas en $\mathcal{C}$. El siguiente ejemplo proporciona un punto de partida.

5.6.6 Ejemplo *Sea* $I : \mathcal{C}([0, 1]), \mathbb{R}) \to \mathbb{R}$ *dada por* $I(f) = \int_0^1 f(x)\, dx$. *Demuéstrese que* I *es continua.*

Solución Debemos mostrar que $f_n \to f$ en $\mathcal{C}$ implica que $I(f_n) \to I(f)$. Pero esto es consecuencia del teorema **5.3.1.** ♦

Ejercicios de §5.6

1. Muéstrese que en el ejemplo **5.6.4**, se puede reemplazar la hipótesis de que f_n está acotada por la de que $f_n(0) = 0$ para obtener la misma conclusión.

2. En **5.6.3**, ¿es la sucesión completa necesariamente convergente?

3. **a.** Muéstrese que el siguiente conjunto es abierto:

$$\left\{ f \in \mathcal{C}([0,1],\mathbb{R}) \,\middle|\, \int_0^1 f(x)\,dx \in\,]0,3[\, \right\}$$

b. Muéstrese que $\mathcal{C}_b$ es cerrado en el espacio de todas las funciones acotadas en un conjunto A.

4. Sea $\mathcal{B} : \mathcal{C}([0, 1]), \mathbb{R})$ un conjunto cerrado, acotado y equicontinuo. Sea $I : \mathcal{B} \to \mathbb{R}$ dada por $I(f) = \int_0^1 f(x)\, dx$. Muéstrese que existe $f_0 \in \mathcal{B}$ para la cual el valor de I es máximo.

5. Sean $f_n : [a, b] \to \mathbb{R}$ funciones continuas, uniformemente acotadas. Sea

$$F_n(x) = \int_a^x f_n(t)\, dt, \quad a \leq x \leq b.$$

Demuéstrese que F_n tiene una subsucesión uniformemente convergente.

§5.7 Principio de la aplicación contractiva y sus aplicaciones

Esta sección trata una importante técnica iterativa en el análisis, llamada el *principio de la aplicación contractiva*. Demuestra resultados de existencia y unicidad, además de dar un procedimiento iterativo específico para localizar la solución. Comenzamos con el método general.

5.7.1 Principio de la aplicación contractiva *Sea M un espacio métrico completo y $\Phi : M \to M$ una transformación dada. Supóngase que existe una constante k, $0 \leq k < 1$ tal que $d(\Phi(x), \Phi(y)) \leq kd(x, y)$ para todo $x, y \in M$. Entonces existe un único punto fijo de Φ; es decir, un punto $x_* \in M$ tal que $\Phi(x_*) = x_*$. De hecho, si x_0 es cualquier punto en M y definimos $x_1 = \Phi(x_0), x_2 = \Phi(x_1), \ldots, x_{n+1} = \Phi(x_n), \ldots$, entonces*

$$\lim_{n\to\infty} x_n = x_*.$$

Intuitivamente, Φ está acortando distancias, de modo que si Φ se itera, los puntos se acumulan, y se acumulan en torno a x_*; véase la figura 5.7-1.

Nota. Existen otros famosos teoremas de punto fijo que tienen una esencia más topológica. Por ejemplo, el teorema del punto fijo de Brouwer (en un caso especial) dice que cualquier función continua $f : D \to D$, donde $D = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$ es el disco unidad, tiene un punto fijo. Es cierto que tales funciones no tienen por qué ser contracciones. Si $n = 1$, esto se demuestra mediante el teorema de valores intermedios (véanse **4.5.5** y el ejercicio **3** de §**4.5**). En este resultado pueden aparecer muchos puntos fijos.

Un resultado relacionado con esto es que cualquier campo vectorial en la superficie de una esfera tiene un cero en alguna parte, o "en cualquier instante, existe un punto sobre la superficie de la Tierra donde el viento no está soplando".

Nuestro teorema de la aplicación contractiva es más analítico, pero también tiene interpretaciones cotidianas: "Tómese un mapa de la ciudad en la que se encuentre y colóquese en una mesa horizontal. Entonces existe un punto en el mapa correspondiente al punto de la mesa directamente debajo del mapa."

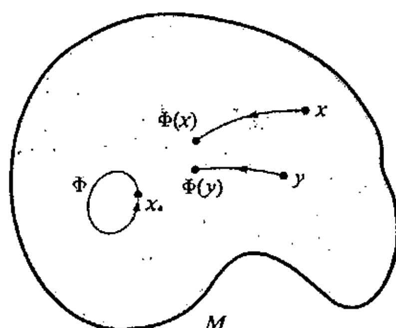


FIGURA 5.7-1 Una contracción acorta las distancias entre puntos

Como primera aplicación, estudiaremos la existencia de soluciones de *ecuaciones diferenciales*. Damos un contexto específico, aunque la técnica es bastante general. La generalizaremos en §**7.5**.

Consideremos una función continua $f(t, x)$ definida en una vecindad de $(t_0, x_0) \in \mathbb{R}^2$. Supongamos que se cumple la siguiente ***condición de Lipschitz***:

$$|f(t, x_1) - f(t, x_2)| \leq K|x_1 - x_2|$$

para todo (t, x_1) y (t, x_2) en una vecindad de (t_0, x_0). Si f es derivable en x y $(\partial f/\partial x)(t, x)$ es continua, entonces la condición se cumple automáticamente (por el teorema del valor medio).

5.7.2 Teorema *Bajo las hipótesis anteriores, existe $\delta > 0$ tal que la ecuación*

$$\frac{dx}{dt} = f(t, x), \quad x(t_0) = x_0 \tag{1}$$

tiene una única solución C^1 $x = \varphi(t)$, tal que $\varphi(t_0) = x_0$, para todo $t_0 - \delta < t < t_0 + \delta$; es decir, $\varphi'(t) = f(t, \varphi(t))$.

La técnica de la demostración se basa en la siguiente observación. Resolver (1) es equivalente a encontrar una función continua φ en $]t_0 - \delta, t_0 + \delta[$ tal que

$$\varphi(t) = x_0 + \int_{t_0}^{t} f(s, \varphi(s))\, ds. \tag{2}$$

La ecuación (2) es una ecuación de punto fijo para la aplicación

$$\Phi : \varphi \mapsto x_0 + \int_{t_0}^{t} f(s, \varphi(s))\, ds$$

que lleva funciones en (un subconjunto adecuado de) C a C. La idea es aplicar el principio de la aplicación contractiva a Φ.

Tenemos, como consecuencia del método usado en el teorema de la aplicación contractiva, un método iterativo específico para resolver (1):

$$\varphi_{n+1} = \Phi(\varphi_n); \quad \text{es decir,} \quad \varphi_{n+1}(t) = x_0 + \int_{t_0}^{t} f(s, \varphi_n(s))\, ds \quad \text{y} \quad \varphi_0 \equiv x_0.$$

Entonces, φ_n converge a la solución deseada (si $|t - t_0|$ es suficientemente pequeño).

La segunda aplicación es a las *ecuaciones integrales*. Comencemos con las llamadas ***ecuaciones de Fredholm***, que tienen la forma

$$f(x) = \lambda \int_a^b K(x, y)\, f(y)\, dy + \varphi(x) \tag{3}$$

donde K y φ son funciones dadas y λ es una constante dada. Supóngase que K y φ son continuas y que para todo x, y tales que $a \leq x \leq b$, $\leq y \leq b$, tenemos

$$|K(x, y)| \leq M. \tag{4}$$

5.7.3 Proposición *Si $\lambda M|b - a| < 1$, entonces la ecuación de Fredholm (3) tiene una única solución.*

La idea esencial de la demostración es muy sencilla. Sea $\Phi : C([a, b], \mathbb{R}) \to C([a, b], \mathbb{R})$ definida por

$$(\Phi(f))(x) = \lambda \int_a^b K(x, y)\, f(y)\, dy + \varphi(x).$$

La aplicación Φ es una contracción, pues

$$d(\Phi(f_1), \Phi(f_2)) = \sup_x \left| \lambda \int_a^b K(x, y)(f_1(y) - f_2(y))\, dy \right| \leq \lambda |b - a| M d(f_1, f_2)$$

y $\lambda |b - a|\, M < 1$, por hipótesis. Así, (3) tiene una única solución, por el teorema de la aplicación contractiva.

Las ***ecuaciones integrales de Volterra*** son de la forma

$$f(x) = \lambda \int_a^x K(x, y)\, f(y)\, dy + \varphi(x). \tag{5}$$

5.7.4 Proposición *Supongamos que K es continua en* $[a, b] \times [a, b]$; *entonces la ecuación integral de Volterra tiene una única solución* $f(x)$ *para* ***cualquier*** λ.

La idea es igual a la anterior, pero ahora tenemos una contracción para $\Phi^N = \Phi \circ \Phi \circ \cdots \circ \Phi$ solamente para N grande; esto compensa el hecho de que no hayamos supuesto que λ es pequeño.

5.7.5 Ejemplo *Encuéntrese un ejemplo de un espacio métrico completo X y una aplicación* $\Phi : X \to X$ *tal que* $d(\Phi(x), \Phi(y)) \leq d(x, y)$ *pero que* Φ ***no*** *tenga un único punto fijo.*

Solución Sea $X = \mathbb{R}$ con la distancia usual $d(x, y) = |x - y|$. Sea $\Phi(x) = x + 1$. Es claro que no existe x tal que $x = x + 1$. Pero $|\Phi(x) - \Phi(y)| = |x - y|$. ♦

Este ejemplo muestra que en el teorema de la aplicación contractiva es esencial tener $k < 1$; $k = 1$ no sirve.

5.7.6 Ejemplo *Muéstrese que el método de aproximaciones sucesivas aplicado a la ecuación diferencial* $f'(x) = f(x)$ *con* $f(0) = 1$ *conduce a la fórmula usual de* e^x.

Solución La ecuación es equivalente a $f(x) = 1 + \int_0^x f(y)\, dy$. Comenzamos con la función cero 0. Puesto que $\Phi(g) = 1 + \int_0^x g(y)\, dy$, obtenemos

$$
\begin{aligned}
\Phi(0)(x) &= 1; \\
\Phi^2(0)(x) &= \Phi(\Phi(0)) = 1 + \int_0^x dy = 1 + x; \\
\Phi(\Phi^2(0))(x) &= 1 + \int_0^x (1 + y)\, dy = 1 + x + \frac{x^2}{2}; \\
\Phi(\Phi^3(0))(x) &= 1 + \int_0^x \left(1 + y + \frac{y^2}{2}\right) dy = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!}; \\
&\cdots \\
\Phi^n(0)(x) &= 1 + x + \cdots + \frac{x^{n-1}}{(n-1)!}.
\end{aligned}
$$

Es claro que esta sucesión converge a e^x. ♦

5.7.7 Ejemplo *Encuéntrese una función f tal que $f'(x) = xf(x)$ para x cercano a 0 y $f(0) = 3$.*

Solución Tomamos como punto de partida la función más sencilla posible que satisfaga la condición inicial, $f_0(x) = 3$, e iteramos el operador de Volterra

$$(Tf)(x) = 3 + \int_0^x tf(t)\,dt$$

en búsqueda de una función f tal que $Tf = f$. Iteramos para producir la siguiente sucesión:

$$\begin{aligned}
f_0(x) &= 3\\
f_1(x) &= (Tf_0)(x) = 3 + \int_0^x t3\,dt = 3 + \frac{3}{2}x^2\\
f_2(x) &= (Tf_1)(x) = 3 + \int_0^x t\left(3 + \frac{3}{2}t^2\right)\,dt = 3 + \frac{3}{2}x^2 + \frac{3}{(2)(4)}x^4\\
f_3(x) &= (Tf_2)(x) = 3 + \int_0^x t\left(3 + \frac{3}{2}t^2 + \frac{3}{(2)(4)}t^4\right)\,dt\\
&= 3 + \frac{3}{2}x^2 + \frac{3}{(2)(4)}x^4 + \frac{3}{(2)(4)(6)}x^6.
\end{aligned}$$

Si continuamos este proceso, vemos que

$$\begin{aligned}
f_n(x) &= 3 + \frac{3}{2}x^2 + \frac{3}{(2)(4)}x^4 + \frac{3}{(2)(4)(6)}x^6 + \cdots + \frac{3}{(2)(4)\cdots(2n)}x^{2n}\\
&= \sum_{k=0}^{n} \frac{3}{(2)(4)\cdots(2k)}x^{2k} = 3\sum_{k=0}^{n} \frac{1}{(2^k)(k!)}x^{2k}\\
&= 3\sum_{k=0}^{n} \frac{1}{k!}\left(\frac{x^2}{2}\right)^k
\end{aligned}$$

Éstas son las sumas parciales de la serie infinita

$$f(x) = 3\sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{x^2}{2}\right)^k,$$

que reconocemos como una serie con un radio infinito de convergencia convergente a la suma

$$f(x) = 3e^{x^2/2}.$$

El cálculo directo de la derivada, mediante la regla de la cadena, confirma que ésta es una solución a nuestro problema (por supuesto, también podríamos encontrarla mediante separación de variables). ♦

5.7.8 Ejemplo *Considérese la ecuación integral*

$$f(x) = a + \int_0^x f(y)xe^{-xy}\,dy.$$

Determínense directamente los intervalos [0, *r*] *en los que se obtiene una contracción.*

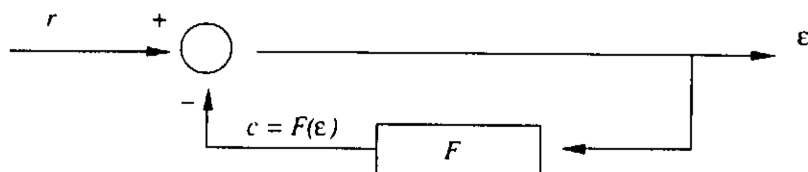
Solución En este caso, $\Phi\,(f(x)) = a + \int_0^x f(y)xe^{-xy}\,dy$, por lo que la constante para Φ es

$$k = \sup_{x\in[0,r]} \int_0^x |K(x,y)|\,dy = \sup_{x\in[0,r]} \int_0^x xe^{-xy}\,dy$$

$$= \sup_{x\in[0,r]} (1 - e^{-x^2}) = 1 - e^{-r^2} < 1.$$

Así, obtenemos una única solución en cualquier intervalo [0, *r*]. ♦

5.7.9 Ejemplo *Consideremos un "ciclo de retroalimentación" que se denota esquemáticamente como sigue:*



En este caso, r, ε, c son funciones de $t \in [a, b]$. Elegimos $M = C([a, b], \mathbb{R})$ y suponemos que $F : M \to M$ es una contracción. Las ecuaciones básicas son $\varepsilon = r - c$; es decir, $\varepsilon + F(\varepsilon) = r$. Demuéstrese la existencia y unicidad de una solución ε de este sistema, dado r.

> **Nota.** Este sistema se puede imaginar como sigue. Una señal (como un impulso eléctrico) r entra al sistema y se une con una señal de retorno c para producir $\varepsilon = r - c$. Una "caja negra" F modifica ε para formar $F(\varepsilon)$; después, la señal c retroalimenta la señal de entrada r.

Solución Obsérvese que para dos soluciones ε_1 y ε_2,

$$\varepsilon_1 + F(\varepsilon_1) = \varepsilon_2 + F(\varepsilon_2)$$

y por lo tanto,

$$F(\varepsilon_1) - F(\varepsilon_2) = \varepsilon_2 - \varepsilon_1, \text{ de modo que } d(\varepsilon_1, \varepsilon_2) = d(F(\varepsilon_1), F(\varepsilon_2)).$$

Como F es una contracción, tenemos una contradicción, a menos que $\varepsilon_1 = \varepsilon_2$. Por lo tanto, si existe una solución ε, ésta es única.

Ahora demostramos la existencia de una solución dada una "entrada" r. Sea

$$G_r(\varepsilon) = r - F(\varepsilon).$$

Queremos que

$$\varepsilon = G_r(\varepsilon).$$

Afirmamos que G_r es una contracción. De hecho,

$$d(G_r(\varepsilon_1), G_r(\varepsilon_2)) = d(r - F(\varepsilon_1), r - F(\varepsilon_2)) = d(F(\varepsilon_1), F(\varepsilon_2)) \leq kd(\varepsilon_1, \varepsilon_2).$$

Como $k < 1$, G_r tiene un punto fijo ε, la única solución del ciclo de retroalimentación. ♦

5.7.10 Ejemplo: un problema de control *Considérese el sistema $dx/dt = f(x) + p$, $x(0) = x_0$, donde $f\colon [0, 1] \to \mathbb{R}$ y $p \in \mathbb{R}$, con f fija. Supongamos que la solución para $p = 0$ es $\bar{x}(t)$, y que y sea cercana a $\bar{x}(1)$. ¿Podemos elegir p de modo que la trayectoria $x(t)$ cumpla $x(0) = x_0$, $x(1) = y$? Es decir, ¿podemos controlar la solución para que termine en y con una elección adecuada de p? Véase la figura 5.7-2. Plantéese este problema como un problema de punto fijo.*

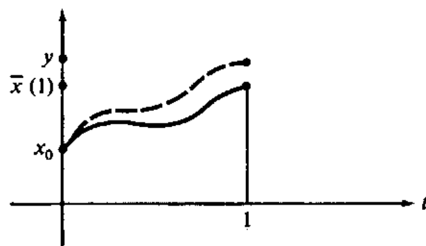


FIGURA 5.7-2 Intento de dirigir una trayectoria hacia el punto y

Solución Mostraremos que esto es posible. Para $p = 0$, tenemos

$$\bar{x}(t) = x_0 + \int_0^t f(\bar{x}(s))\, ds$$

y en consecuencia,

$$\bar{x}(1) = x_0 + \int_0^1 f(\bar{x}(s))\, ds.$$

Queremos p tal que $y = x_0 + \int_0^1 f(x(s))\, ds + p$, donde $x(s)$, que depende de p, satisface $dx/dt = f + p$. Así,

$$\left.\begin{aligned} p &= y - x_0 - \int_0^1 f(x(s))\, ds, \\ x(t) &= x_0 + \int_0^t f(x(s))\, ds + pt. \end{aligned}\right\} \qquad (6)$$

Sea $\Phi : \mathbb{R} \times C([0, 1]) \to \mathbb{R} \times C([0, 1])$ dada por $\Phi : (p, x) \to$ miembro derecho de (6). Si Φ es una contracción, el punto fijo de Φ será la solución del problema de control. ♦

Ejercicios de §5.7

1. ¿Para qué valores de α es $\Phi(x) = \alpha x$ una contracción en $\mathbb{R}$?

2. Encuéntrese un desarrollo en serie en el intervalo $[0, \frac{1}{2}]$ para la solución de la ecuación (1) si $K(x, y) = x$ y $a = 1$.

3. ¿Para qué intervalos $[0, r]$, $r \leq 1$ es $f : [0, r] \to [0, r]$, $x \mapsto x^2$ una contracción?

4. Muéstrese que el sistema de ecuaciones

$$x_1 = \frac{1}{4}x_1 - \frac{1}{4}x_2 + \frac{2}{15}x_3 + 3$$

$$x_2 = \frac{1}{4}x_1 + \frac{1}{5}x_2 + \frac{1}{2}x_3 - 1$$

$$x_3 = -\frac{1}{4}x_1 + \frac{1}{3}x_2 - \frac{1}{3}x_3 + 2$$

tiene una única solución, usando el principio de la aplicación contractiva. [*Sugerencia*: se puede elegir una norma adecuada en $\mathbb{R}^3$, o bien hacer una estimación mediante la desigualdad de Schwarz.]

5. Conviértase $dy/dx = 3xy$, $y(0) = 1$ en una ecuación integral y plantéese un esquema iterativo para resolverla.

6. Plantéese un esquema iterativo para resolver $dx/dt = tx^2 + x^3$, $x(0) = 1$. Tómese $\delta < 1/5$.

7. Considérese $dx/dt = x^2$, $x(0) = 1$ y resuélvase de forma explícita para ver que δ en **5.7.2** es finita.

8. Sea M un espacio métrico compacto y $\Phi: M \to M$ tal que $d(\Phi(x), \Phi(y)) < d(x, y)$ para todo $x, y \in M$, $x \neq y$.

a. Muéstrese que Φ tiene un único punto fijo. [*Sugerencia:* minimícese $d(\Phi(x), x)$.]

b. Muéstrese que **a** es falso si M no es compacto (encuéntrese un contraejemplo).

§5.8 Teorema de Stone-Weierstrass

En el estudio de las funciones continuas y la convergencia uniforme, tres de los resultados más básicos son el teorema de Arzela-Ascoli, el principio de la aplicación contractiva discutido en §**5.7** y el teorema de Stone-Weierstrass, que analizaremos ahora.

El propósito del teorema de Stone-Weierstrass es mostrar que cualquier función continua se puede aproximar de manera uniforme mediante una función que tenga propiedades más manejables, como un polinomio. Tales técnicas de aproximación polinómica son importantes tanto teóricamente como en el aspecto numérico.

Comenzaremos dando el resultado para el caso particular del intervalo unidad. Éste fue demostrado por primera vez por Weierstrass, pero aquí presentaremos una versión de Sergei Bernstein (1880-1968).[2]

5.8.1 Teorema *Sean $f: [0, 1] \to \mathbb{R}$ continua y $\varepsilon > 0$. Entonces existe un polinomio $p(x)$ tal que $\|p - f\| < \varepsilon$. De hecho, la sucesión de* ***polinomios de Bernstein***

$$p_n(x) = \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1 - x)^{n-k}$$

converge uniformemente a f si $n \to \infty$, donde

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

denota el coeficiente binómico.

[2] Bernstein fue un matemático ruso muy influyente en el análisis y la teoría de la probabilidad. La demostración está en su trabajo *Démonstration du théorème de Weierstrass, fondée sur le calcul des probabilités*. Commun. Soc. Math. Kharkov (2), **13** (1912–13), 1–2.

El primer enunciado del teorema es una consecuencia del segundo; este último se puede entender fácilmente si uno sabe algo de teoría de probabilidad, lo cual se asume sólo en lo que respecta al párrafo que sigue en nuestra exposición. Huelga decir que el conocimiento de Bernstein de la teoría de probabilidad le ayudó, sin duda, a comprender y demostrar este teorema.

Imaginemos una "moneda" con probabilidad x de caer con la cara hacia arriba y, en consecuencia, con probabilidad $1 - x$ de caer con la cruz hacia arriba. En n lanzamientos, la probabilidad de obtener exactamente k caras es

$$\binom{n}{k} x^k(1 - x)^{n-k}.$$

Supóngase que en un juego "de cara y cruz", se pagan $f(k/n)$ monedas cuando se obtienen exactamente k caras al tirar la moneda n veces. La cantidad promedio (después de jugar mucho tiempo) que se paga al jugar n veces es

$$\sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k(1 - x)^{n-k} = p_n(x).$$

En este caso, $f(k/n)$ es f para la fracción de tiradas que salen cara. Ahora imaginemos que n es muy grande. Entonces esperamos que al jugar n veces, k/n esté muy cercano a x, que es la probabilidad de que salgan k caras (igual a la fracción de tiempo en que salen las k caras), de modo que nuestro pago promedio será muy cercano a $f(x)$. Por lo tanto, si n es grande, esperamos que $p_n(x)$ sea cercano a $f(x)$. Ésta es la razón intuitiva de la validez de este resultado. La demostración real es un poco complicada, como podría esperarse de la complejidad del juego.

Aun para una función f sencilla como $f(x) = \sqrt{x}$, no es trivial encontrar un polinomio que la aproxime.

Podemos volver a enunciar el teorema como sigue. Sea $\mathcal{P}$ el conjunto de todos los polinomios $p: [0, 1] \to \mathbb{R}$. Entonces, la primera afirmación establece que $\mathcal{P}$ es *denso* en $C([0, 1], \mathbb{R})$; es decir, que $\text{cl}(\mathcal{P}) = C([0, 1], \mathbb{R})$.

En 1937, Marshall Stone descubrió una generalización muy útil de este teorema, permitiendo el uso de conjuntos más generales que $[0, 1]$ y reemplazando $\mathcal{P}$ por una familia general de funciones que satisfacen ciertas propiedades. La demostración hace uso del caso particular ya demostrado. El teorema es útil en muchas ramas del análisis (por ejemplo, lo usaremos en nuestro estudio del análisis de Fourier).

5.8.2 Teorema de Stone-Weierstrass *Sean M un espacio métrico, $A \subset M$ un conjunto compacto y $\mathcal{B} \subset C(A, \mathbb{R})$, con las siguientes condiciones:*

i. *$\mathcal{B}$ es un **álgebra**; si $f, g \in \mathcal{B}$ y $\alpha \in \mathbb{R}$, entonces $f + g \in \mathcal{B}$, $f \cdot g \in \mathcal{B}$, y $\alpha f \in \mathcal{B}$*

ii. *La función constante $x \mapsto 1$ está en $\mathcal{B}$.*

iii. *$\mathcal{B}$ separa puntos; es decir, para $x, y \in A$, $x \neq y$ existe $f \in \mathcal{B}$ tal que $f(x) \neq f(y)$.*

Entonces $\mathcal{B}$ es denso en $C(A, \mathbb{R})$; es decir, $\text{cl}(\mathcal{B}) = C(A, \mathbb{R})$.

5.8.3 Ejemplo *Sea p_n una sucesión uniformemente convergente de polinomios en $[0, 1]$ y $f = \lim_{n\to\infty} p_n$. ¿Debe f ser derivable?*

Solución No; por **5.8.1**, cualquier función continua es límite de polinomios y existen muchas funciones continuas que no son derivables, como, por ejemplo

$$f(x) = \begin{cases} 0, & 0 \leq x \leq 1/2, \\ 2x - 1, & 1/2 \leq x \leq 1. \end{cases} \quad ♦$$

5.8.4 Ejemplo *Demuéstrese* (**a**) *directamente de* **5.8.1** *y* (**b**) *a partir de* **5.8.2**, *que los polinomios en $[a, b]$ son densos en $C([a, b], \mathbb{R})$.*

Solución

a. Por **5.8.1**, si $f: [0, 1] \to \mathbb{R}$ es continua y $\varepsilon > 0$, entonces existe un polinomio p tal que $\|f - p\| < \varepsilon$. Sea $g : [a, b] \to \mathbb{R}$ continua. Hacemos un cambio de escala y definimos

$$f(x) = g(x(b - a) + a), \quad 0 \leq x \leq 1,$$

de modo que $f: [0, 1] \to \mathbb{R}$. Sea

$$q(x) = p\left(\frac{x - a}{b - a}\right), \quad a \leq x \leq b,$$

de modo que $q : [a, b] \to \mathbb{R}$. Así, q también es un polinomio. Afirmamos que $\|g - q\| < \varepsilon$. De hecho,

$$|g(x) - q(x)| = \left| f\left(\frac{x - a}{b - a}\right) - p\left(\frac{x - a}{b - a}\right) \right|,$$

y entonces $\|g - q\| < \varepsilon$, pues $\|f - p\| < \varepsilon$. Así, los polinomios en $[a, b]$ son densos.

b. En el teorema **5.8.2**, sean $A = [a, b]$ y $\mathcal{B} = \{q \in C([a, b], \mathbb{R}) \mid q \text{ es un polinomio}\}$. Entonces $\mathcal{B}$ satisface **i** y **ii**. También satisface **iii**, ya que si $x \neq y$, podemos hacer

$$f(t) = t$$

de modo que $f(x) \neq f(y)$. Así, $\mathcal{B}$ es denso, por el teorema **5.8.2**. ♦

Ejercicios de §5.8

1. Muéstrese que existe un polinomio $p(x)$ tal que $|p(x) - \text{sen}\, x| < 1/100$ para $0 \leq x \leq 2\pi$.

2. Supóngase que p_n es una sucesión de polinomios que convergen uniformemente a f en $[0, 1]$ y que f *no* es un polinomio. Demuéstrese que los grados de p_n no están acotados. [*Sugerencia*: un polinomio p de grado N queda determinado de forma única mediante sus valores en $N + 1$ puntos $x_0, \ldots, x_N$ por medio de la ***fórmula de interpolación de Lagrange***

$$p(x) = \sum_{i=0}^{N} \pi_i(x) \frac{p(x_i)}{\pi_i(x_i)},$$

donde $\pi_i(x) = (x - x_0)(x - x_1) \cdots (x - x_N)/(x - x_i)$.]

3. Demuéstrese que el conjunto de los polinomios en $C([a, b], \mathbb{R})$ no es abierto. ¿Puede un subconjunto de un espacio métrico ser abierto y denso *a la vez*?

4. Considérese el conjunto de todos los polinomios $p(x, y)$ de dos variables $x, y \in [0, 1] \times [0, 1]$. Demuéstrese que este conjunto es denso en $C([0, 1] \times [0, 1], \mathbb{R})$.

5. Considérese el conjunto de todas las funciones en $[0, 1]$ de la forma

$$h(x) = \sum_{j=1}^{n} a_j e^{b_j x}, \quad \text{donde } a_j,\ b_j \in \mathbb{R}.$$

¿Es este conjunto denso en $C([0, 1], \mathbb{R})$?

6. Si $T = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$, entonces una función $f\colon T \to \mathbb{R}$ se puede imaginar como una función periódica de $\mathbb{R}$ en $\mathbb{R}$ de periodo 2π, identificando previamente un número θ con el punto $(\cos\theta, \text{sen}\,\theta)$ en T

$$\theta \to (\cos\theta, \text{sen}\,\theta) \xrightarrow{f} f(\theta).$$

Análogamente, cualquier función periódica de periodo 2π en $\mathbb{R}$ se puede concebir como una función en T. Una suma finita de senos y cosenos de la forma

$$\frac{a_0}{2} + \sum_{k=1}^{N} (a_k \cos(k\theta) + b_k \text{sen}(k\theta))$$

se denomina ***polinomio trigonométrico*** (N es arbitrario pero finito).

Muéstrese que si f es una función periódica de $\mathbb{R}$ en $\mathbb{R}$ de periodo 2π y $\varepsilon > 0$, entonces existe un polinomio trigonométrico q tal que $|q(\theta) - f(\theta)| < \varepsilon$ para todo θ en $\mathbb{R}$ (el número de términos N puede depender de f y de ε).

§5.9 Criterios de Abel y Dirichlet

El criterio M de Weierstrass falla en ciertos casos en los que queremos determinar si tenemos convergencia uniforme; para esos casos, los matemáticos han diseñado otros criterios. El primer criterio que mostraremos en esta sección fue creado por el matemático noruego Niels Abel, mientras que el segundo se atribuye a P.G. Dirichlet, alemán de origen francés; ambos trabajaron durante la primera mitad del siglo XIX. Estos criterios son útiles en muchos ejemplos y son de particular utilidad para el estudio de las series de potencias y de las series de Fourier. Son importantes cuando tenemos convergencia uniforme pero no absoluta.

5.9.1 Criterio de Abel *Sean $A \subset \mathbb{R}^m$ y $\varphi_n : A \to \mathbb{R}$ una sucesión decreciente de funciones; es decir, $\varphi_{n+1}(x) \leq \varphi_n(x)$ para cada $x \in A$. Supóngase que existe una constante M tal que $\varphi_n(x) \,|\leq M$ para todo $x \in A$ y todo n. Si $\sum_{n=1}^{\infty} f_n(x)$ converge uniformemente en A, entonces también lo hace $\sum_{n=1}^{\infty} \varphi_n(x) f_n(x)$.*

Obtenemos algunos criterios útiles para las series ordinarias si tomamos el caso particular en que φ_n y f_n son funciones constantes. También existe un criterio similar cuando φ_n son crecientes, el cual se puede deducir aplicando el criterio de Abel a $-\varphi_n$. Un criterio relacionado con éste es el de Dirichlet.

5.9.2 Criterio de Dirichlet *Sea $s_n(x) = \sum_{i=1}^{n} f_i(x)$ para una sucesión $f_n : A \subset \mathbb{R}^m \to \mathbb{R}$. Supóngase que existe M constante tal que $|s_n(x)| \leq M$ para todo $x \in A$ y todo n. Sea $g_n : A \subset \mathbb{R}^m \to \mathbb{R}$ tal que $g_n \to 0$ uniformemente, $g_n \geq 0$ y $g_{n+1}(x) \leq g_n(x)$. Entonces $\sum_{n=1}^{\infty} f_n(x) g_n(x)$ converge uniformemente en A.*

Por ejemplo, considérese la serie alternante $\sum(-1)^n g_n(x)$, donde $g_n \geq 0$, $g_n(x) \to 0$ uniformemente y $g_{n+1} \leq g_n$. Sea $f_n(x) = (-1)^n$. Entonces $|s_n(x)| \leq 1$, de modo que $\sum(-1)^n g_n(x)$ converge uniformemente. Obsérvese que, como un caso particular del teorema **5.9.2**, una serie alternante cuyos términos decrecen a cero es convergente.

Obsérvense las analogías y las diferencias entre estos teoremas. Las condiciones sobre φ_n en el criterio de Abel no implican que φ_n converja uniformemente. Además, no pedimos que $\varphi_n \geq 0$. Las demostraciones de estos teoremas utilizan una relación llamada *fórmula de sumación parcial de Abel.*

5.9.3 Ejemplo *Muéstrese que $\sum_1^{\infty} (\text{sen } nx)/n$ converge uniformemente en $[\delta, 2\pi - \delta]$, donde $0 < \delta < 2\pi$.*

Solución Aplicamos el criterio de Dirichlet con $f_n(x) = \text{sen } nx$ y $g_n(x) = 1/n$. La única hipótesis que no es obvia es $\sum_{l=1}^{n} f_l(x)| \leq M$. Para mostrarla, necesitamos una técnica especial; usaremos la identidad trigonométrica

$$2\,\text{sen}(lx)\text{sen}\left(\frac{1}{2}x\right) = \cos\left[\left(l - \frac{1}{2}\right)x\right] - \cos\left[\left(l + \frac{1}{2}\right)x\right]$$

y sumamos de $l = 1$ a $l = n$. Obtenemos una suma telescópica, de modo que

$$\left|2\,\text{sen}\left(\frac{1}{2}x\right)(\text{sen } x + \cdots + \text{sen } nx)\right| = \left|\cos\left(\frac{1}{2}x\right) - \cos\left(n + \frac{1}{2}x\right)\right| \leq 2.$$

Así,

$$|\text{sen } x + \cdots + \text{sen } nx| \leq \frac{1}{\left|\text{sen}\frac{1}{2}x\right|},$$

lo que da una cota para $\sum_{l=1}^{n} f_l(x)$, la cual sirve mientras $\text{sen}(x/2)$ se mantenga lejos de cero. Por ejemplo, en $[\delta, 2\pi - \delta]$ obtenemos dicha cota. Obsérvese que los argumentos necesarios en este caso son más finos que el criterio M. ♦

5.9.4 Ejemplo *Muéstrese que* $\sum_{n=1}^{\infty} \frac{(-1)^n}{n} e^{-nx}$ *converge uniformemente en* $[0, \infty[$.

Solución Esta vez aplicamos el criterio de Abel. Sea $\varphi_n(x) = e^{-nx}$. Para $x \geq 0$, φ_n es decreciente y $|e^{-nx}| \leq 1$ (¿por qué?). Ya sabemos que la serie alternante $\sum_1^{\infty}(-1)^n/n$ converge y entonces, por el criterio de Abel, la serie converge uniformemente. ♦

5.9.5 Ejemplo *Sea* $f(x) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n} e^{-nx}$. *Muéstrese que f es continua en* $[0, \infty[$.

Solución La solución es inmediata del ejemplo anterior y del hecho de que el límite uniforme de funciones continuas es continuo. ♦

Ejercicios de §5.9

Verifíquese la convergencia o convergencia uniforme de las siguientes series:

1. $\sum_{1}^{\infty} \frac{x^n}{n!} e^{-nx}, \quad 0 \leq x \leq 1.$

2. $\sum_{1}^{\infty} \frac{(-1)^n x^n}{n}, \quad 0 \leq x \leq 1.$

3. $\sum_{1}^{\infty} \frac{(-1)^n}{n+x}, \quad 0 \leq x \leq \infty.$

4. $\sum_{1}^{\infty} \frac{\operatorname{sen} nx}{n} e^{-nx}, \quad 0 \leq \delta \leq x \leq \pi - \delta.$

5. $\sum_{1}^{\infty} nx^n, \quad 0 \leq x < 1.$

§5.10 Series de potencias y sumabilidad Cesaro y Abel

En esta sección estudiaremos algunos temas adicionales y opcionales de la teoría de series infinitas. Comenzaremos con las series de potencias.

5.10.1 Definición *Una serie de potencias es una serie de la forma*

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k$$

donde los coeficientes a_k son números reales (o complejos) fijos. Sea

$$\operatorname{lím\,sup}_{k \to \infty} \sqrt[k]{|a_k|} = \frac{1}{R};$$

R es el ***radio de convergencia*** *de la serie de potencias y el conjunto $\{x \mid |x - x_0| = R\}$ es el* ***círculo de convergencia*** *(x puede ser real o complejo).*

Obsérvese que $0 \leq R \leq +\infty$; R puede ser 0 o $+\infty$. La razón de la terminología en la definición **5.10.1** se aclara en el siguiente resultado.

5.10.2 Teorema *$\sum_{k=0}^{\infty} a_k(x-x_0)^k$ converge absolutamente para $|x-x_0|<R$, converge uniformemente para $|x-x_0| \leq R'$, donde $R' < R$ y diverge si $|x-x_0| > R$ (el teorema no proporciona información si $|x-x_0| = R$).*

Está claro que estas propiedades de convergencia definen R de forma única.

5.10.3 Corolario *La suma de una serie de potencias es una función C^{∞} dentro de su círculo de convergencia. Se puede derivar término a término; la serie derivada tiene el mismo radio de convergencia.*

La demostración usa los resultados de §**5.3** acerca de la derivación término a término de las series. Si $\lim_{n\to\infty} |a_n/a_{n+1}|$ existe, entonces el radio de convergencia es

$$R = \lim_{n\to\infty} \left|\frac{a_n}{a_{n+1}}\right|.$$

Esto se ve del teorema **5.10.2** junto con el criterio del cociente. Pedimos al lector que lo demuestre.

A continuación, analizamos el concepto de ***sumabilidad Cesaro***.

5.10.4 Definición *Sean $S_n = \sum_{k=1}^{n} a_k$ y $\sigma_n = (\sum_{k=1}^{n} S_k)/n$; entonces, σ_n es la media aritmética de las primeras n sumas parciales de la serie. La serie $\sum_{k=1}^{\infty} a_k$ es* ***sumable Cesaro de orden 1*** *o sumable $(C, 1)$ a A si $\lim_{n\to\infty} \sigma_n = A$. En este caso, escribimos*

$$\sum_{k=1}^{\infty} a_k = A \ (C, 1).$$

La idea es encontrar una forma de dar significado a series que en otro caso serían divergentes. Por ejemplo,

$$1 - 1 + 1 - 1 + \cdots = \frac{1}{2} \ (C, 1).$$

En este caso, las sumas parciales son $S_n = 1, 0, 1, 0, \ldots$, y sus promedios son

$$\sigma_n = \frac{1}{n} \sum_{k=1}^{n} S_k = \frac{1}{1}, \frac{1}{2}, \frac{2}{3}, \frac{2}{4}, \frac{3}{5}, \frac{3}{6}, \ldots.$$

Así, $\sigma_{2n} = n/2n$, $\sigma_{2n+1} = (n+1)/(2n+1)$ y entonces $\lim_{n\to\infty} \sigma_n = 1/2$. Obsérvese la fórmula general

$$\sigma_n = \sum_{k=1}^{n} \left(1 - \frac{k-1}{n}\right) a_k.$$

Sin embargo, podemos introducir un método *aún más poderoso* de sumación al promediar las σ_n, al igual que el método $(C, 1)$ se basó en los promedios de las S_n. Es decir, definimos la suma $(C, 2)$ de la serie dada como $\text{lím}_{n\to\infty} \sigma_1 + \sigma_2 + \cdots + \sigma_n)/n$, si el límite existe.

El lector puede ver ahora con facilidad cómo se definiría la sumabilidad (C, r) para valores arbitrarios $r = 1, 2, \ldots$, con lo que se obtendrían sucesivamente métodos más poderosos de sumación. A continuación presentamos algunas propiedades básicas de la sumabilidad $(C, 1)$.

i. Si $\sum_{k=1}^{n} a_k = A\ (C, 1)$ y $\sum_{k=1}^{n} b_k = B\ (C, 1)$, entonces

$$\sum_{k=1}^{\infty}(\alpha a_k + \beta b_k) = \alpha A + \beta B\ (C, 1).$$

ii. Si $\sum_{k=1}^{\infty} a_k = A\ (C, 1)$, entonces

$$\sum_{k=1}^{\infty} a_{k+1} = A - a_1\ (C, 1) \quad (\textit{decapitación}).$$

iii. Si $\sum_{k=1}^{\infty} a_k = A$ en el sentido usual, entonces

$$\sum_{k=1}^{\infty} a_k = A\ (C, 1) \quad (\textit{regularidad}).$$

(Esta propiedad es crucial; cualquier método respetable de sumación debe poseerla.)

Demostración

iii. Tenemos que $S_n \to A$. Entonces, dado $B < A$, existe n_0 tal que $n \geq n_0$ implica $S_n \geq B$. Entonces,

$$\sigma_n = \frac{1}{n}(S_1 + \cdots + S_{n_0} + S_{n_0+1} + \cdots + S_n) \geq \frac{1}{n}(S_1 + \cdots + S_{n_0}) + \frac{n - n_0}{n}B.$$

Por lo tanto, $\text{lím inf}_{n\to\infty} \sigma_n \geq B$: Como $B < A$ era arbitrario, $\text{lím inf}_{n\to\infty} \sigma_n \geq A$. Un argumento similar muestra que $\text{lím sup}_{n\to\infty} \sigma_n \leq A$. Así,

$$A \leq \text{lím} \inf_{n\to\infty} \sigma_n \leq \text{lím} \sup_{n\to\infty} \sigma_n \leq A.$$

De modo que ambos límites son iguales, y por lo tanto $\text{lím}_{n\to\infty} \sigma_n = A$.

A continuación analizamos otro método de sumación llamado ***sumación de Abel*** (aunque en realidad fue ideado por Euler).

5.10.5 Definición *$\sum_{k=0}^{\infty} a_k$ es **sumable a A en el sentido de Abel** si* $\lim_{x\to 1^-} \sum_{k=0}^{\infty} a_k x^k = A$. *Escribimos* $\sum_{k=0}^{\infty} a_k = A$ *(Abel).*

Por ejemplo, de nuevo tenemos que

$$\frac{1}{2} = 1 - 1 + 1 - 1 + \cdots \quad \text{(Abel)}$$

pues $f(x) = 1 - x + x^2 - \cdots = 1/(1 + x)$ para $|x| < 1$ y esto tiende a $1/2$ cuando $x \to 1^-$.

Obsérvese que (al menos en este ejemplo) el método de Abel da el mismo resultado que el método $(C, 1)$. En realidad, esto siempre ocurre, como veremos en el teorema **5.10.7**. Primero demostraremos que la sumabilidad de Abel es *regular*.

5.10.6 Teorema (Abel) *Si* $\sum_{k=0}^{\infty} a_k = A$, *entonces* $\sum_{k=0}^{\infty} a_k x^k$ *converge para* $|x| < 1$ *y* $\lim_{x\to 1^-} \sum_{k=0}^{\infty} a_k x^k = A$.

Así, si una serie de potencias converge en un intervalo *cerrado*, su suma es continua, incluso en los extremos.

En realidad, el método de Abel es más poderoso que el método $(C, 1)$.

5.10.7 Teorema $\sum_{k=0}^{\infty} a_k = A\ (C, 1)$ *implica* $\sum_{k=0}^{\infty} a_k = A$ *(Abel).*

Es interesante buscar condiciones bajo las cuales una serie sumable Cesaro (o sumable Abel, etcétera) en realidad es convergente en el sentido usual. Daremos ahora un resultado de G.H. Hardy.

5.10.8 Teorema *Si* $\sum a_n = A\ (C, 1)$ *y* $a_n = O(1/n)$ *(es decir, si* $|a_n| \le C/n$ *para C constante), entonces* $\sum a_n$ *converge (a A) en el sentido usual.*

Nota. Los teoremas como el anterior se conocen como "tauberianos", en honor de A. Tauber, quien demostró un teorema que relaciona la sumabilidad de Abel con la convergencia ordinaria.

5.10.9 Ejemplo *Determínese el radio de convergencia de $\sum_{k=0}^{\infty} x^k$ y de $\sum_{k=0}^{\infty} x^k/k!$*

Solución En estos casos podemos usar la fórmula

$$R = \lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right|.$$

El primer ejemplo da como resultado $R = 1$, y el segundo da

$$R = \lim_{n\to\infty} \left(\frac{(n+1)!}{n!} \right) = \lim_{n\to\infty} (n+1) = \infty.$$

Así, $\sum_{k=0}^{\infty} x^k$ converge a $1/(1-x)$ si $|x| < 1$ y $\sum_{k=0}^{\infty} x^k/k!$ converge a e^x para todo x. ♦

5.10.10 Ejemplo *Muéstrese que $\sum_{k=1}^{\infty} (-1)^k k$ no es sumable* (C, 1).

Solución En este caso,

$$\begin{aligned} &a_n : -1, +2, -3, +4, -5, +6, \ldots \\ &S_n : -1, +1, -2, +2, -3, +3, \ldots \\ &T_n : -1, 0, -2, 0, -3, 0, \ldots \\ &\sigma_{2n} = 0,\ \sigma_{2n-1} = -n/(2n-1) \to -1/2. \end{aligned}$$

Así, $\lim_{n\to\infty} \sigma_n$ *no* existe. Sin embargo, la suma (C, 2) es $-1/4$ (el estudiante debe verificar esta afirmación). ♦

5.10.11 Ejemplo *Muéstrese que $\sum_{k=1}^{\infty} (-1)^k k = -1/4$ (Abel). En este caso,*

$$\sum_{k=1}^{\infty} (-1)^k k x^k = x \frac{d}{dx} \sum_{k=1}^{\infty} (-1)^k x^k = x \frac{d}{dx} \frac{1}{1+x} = -\frac{x}{(1+x)^2}, \quad |x| < 1.$$

Esto tiende a $-1/4$ cuando $x \to 1^-$, de modo que

$$\sum_{k=1}^{\infty} (-1)^k k = -\frac{1}{4} \text{ (Abel)}.$$

Ejercicios de §5.10

1. Calcúlese el radio de convergencia de

$$\sum_{k=1}^{\infty} \frac{x^k}{k^2} \quad \text{y de} \quad \sum_{k=1}^{\infty} k!x^k.$$

2. Muéstrese que

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} kx^{k-1} = \sum_{k=0}^{\infty}(k+1)x^k, \quad -1 < x < 1$$

derivando una serie adecuada.

3. Muéstrese que

$$\frac{2}{3} = 1 + 0 - 1 + 1 + 0 - 1 + 1 + 0 - 1 + \cdots \quad (C, 1).$$

(Obsérvese que la inserción de ceros puede alterar la suma Cesaro.)

4. Muéstrese que

$$1 + 0 - 1 + 1 + 0 - \cdots = \frac{2}{3}$$

(Abel).

5. Muéstrese que

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \frac{\pi}{4}$$

obteniendo un desarrollo en serie para el arco tangente y usando el teorema de Abel.

Demostraciones de los teoremas del capítulo 5

5.1.4 Proposición *Sean $A \subset M$, donde M es un espacio métrico, y $f_k : A \to N$ una sucesión de funciones continuas tales que $f_k \to f$ (uniformemente en A). Entonces f es continua en A.*

Demostración Como $f_n \to f$ uniformemente, dado $\varepsilon > 0$ podemos encontrar un entero N tal que $k \geq N$ implique $\rho(f_k(x), f(x)) < \varepsilon/3$ para todo $x \in A$. Consideremos un punto particular $x_0 \in A$. Como f_N es continua, existe $\delta > 0$ tal que $d(x, x_0) < \delta, x \in A \Rightarrow \rho(f_N(x), f_N(x_0)) < \varepsilon/3$. Entonces, para $d(x, x_0) < \delta, \rho(f(x), f(x_0)) \leq \rho(f(x), f_N(x_0)) + \rho(f_N(x), f_N(x_0)) + \rho(f_N(x_0), f(x_0)) < \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon$. Como x_0 es arbitrario, f es continua en cada punto de A; por lo tanto, es continua. ■

5.2.1 Criterio de Cauchy *Sea N un espacio métrico con métrica ρ y sea A un conjunto. Supóngase que N es completo y que $f_k : A \to N$ es una sucesión de funciones. Entonces f_k converge uniformemente en A sii para cada $\varepsilon > 0$ existe N tal que $l, k \geq N$ implica $\rho(f_k(x), f_l(x)) < \varepsilon$ para todo $x \in A$.*

Demostración Si $f_k \to f$ uniformemente, entonces dado $\varepsilon > 0$, podemos determinar un entero N tal que $k \geq N$ implique $\rho(f_k(x), f(x)) < \varepsilon/2$ para todo x. Entonces, si k, $l \geq N$, $\rho(f_k(x), f_l(x)) \leq \rho(f_k(x), f(x)) + \rho(f(x), f_l(x)) < \varepsilon/2 + \varepsilon/2 = \varepsilon$.

Recíprocamente, si, dado que $\varepsilon > 0$ podemos encontrar un N tal que $k, l \geq N$ implique $\rho(f_k(x), f_l(x)) < \varepsilon$ para todo x, por lo que $f_k(x)$ es una sucesión de Cauchy en cada punto x, y entonces $f_k(x)$ converge puntualmente a algo, que denotamos $f(x)$. Además, podemos encontrar N tal que $k, l \geq N$ implique $\rho(f_k(x), f_l(x)) < \varepsilon/2$ para todo x. Como $f_k(x) \to f(x)$ en cada punto x, podemos determinar para cada x un N_x tal que $l \geq N_x$ implique $\rho(f_l(x), f(x)) < \varepsilon/2$. Sea $l \geq \text{máx}\{N, N_x\}$. Entonces $k \geq N$ implica $\rho(f_k(x), f(x)) \leq \rho(f_k(x), f_l(x)) + \rho(f_l(x), f(x)) < \varepsilon/2 + \varepsilon/2 = \varepsilon$. Como esto es cierto para cada punto x, hemos encontrado N tal que $k \geq N$ implique $\rho(f_k(x), f(x)) < \varepsilon$ para todo x. Por lo tanto, $f_k \to f$ (uniformemente). ■

5.2.2 Criterio M de Weierstrass *Supongamos que V es un espacio vectorial normado completo y que $g_k : A \to V$ son funciones tales que existen constantes M_k que cumplen $\|g_k(x)\| \leq M_k$ para cada $x \in A$ y $\sum_{k=1}^{\infty} M_k$ converge. Entonces $\sum_{k=1}^{\infty} g_k$ converge uniformemente (y absolutamente).*

Demostración Como $\sum_{k=1}^{\infty} M_k$ converge, para todo $\varepsilon > 0$ existe N tal que $k \geq N$ implique $|M_k + \cdots + M_{k+p}| < \varepsilon$ para todo $p = 1, 2, \ldots$. Para $k \geq N$, la desigualdad triangular implica

$$\|g_k(x) + \cdots + g_{k+p}(x)\| \leq \|g_k(x)\| + \cdots + \|g_{k+p}(x)\| \leq M_k + \cdots + M_{k+p} < \varepsilon$$

para todo $x \in A$. Así, por el criterio de Cauchy para series, $\sum_{k=1}^{\infty} g_k$ converge uniformemente. ■

5.3.1 Teorema *Supongamos que $f_1, f_2, f_3, \ldots$ son funciones integrables en un intervalo cerrado y acotado $[a, b]$ que convergen uniformemente a una función límite f en $[a, b]$. Entonces f es integrable en $[a, b]$* y

$$\lim_{n \to \infty} \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx.$$

Demostración Supongamos, en primer lugar, que f es integrable. Para demostrar la relación del límite, recuérdese que si $|g(x)| \leq M$ entonces

$$\left|\int_a^b g(x)\,dx\right| \leq M(b-a).$$

Dado que $\varepsilon > 0$, sea N tal que $k \geq N$ implique $|f_k(x) - f(x)| < \varepsilon/(b-a)$. Entonces

$$\left|\int_a^b f_k(x)\,dx - \int_a^b f(x)\,dx\right| = \left|\int_a^b (f_k(x) - f(x))\,dx\right| \leq \varepsilon \cdot \frac{b-a}{b-a} = \varepsilon$$

como se pedía.

La parte difícil es mostrar que f es integrable. Para esto, usamos la definición **4.8.1**, que es equivalente a

$$\underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx.$$

Sea $\varepsilon > 0$ dada y sea N tal que $|f_n(x) - f(x)| < \varepsilon$ si $n \geq N$ y $a \leq x \leq b$. En particular, esto implica que f está acotada (¿por qué?). Obsérvese que si α es una cota superior de f_n en $[x_i, x_{i+1}]$, entonces $\alpha + \varepsilon$ es una cota superior de f en el mismo intervalo. Así,

$$\sup\{f(x) \mid x \in [x_i, x_{i+1}]\} \leq \sup\{f_n(x) \mid x \in [x_i, x_{i+1}]\} + \varepsilon$$

y, por lo tanto,

$$U(f, P) \leq U(f_n, P) + \varepsilon(b-a).$$

Análogamente,

$$L(f_n, P) - \varepsilon(b-a) \leq L(f, P).$$

Al tomar los ínfimos y supremos sobre P, obtenemos

$$\overline{\int_a^b} f(x)\,dx \leq \int_a^b f_n(x)\,dx + \varepsilon(b-a)$$

y

$$\int_a^b f_n(x)\,dx - \varepsilon(b-a) \leq \overline{\int_a^b} f(x)\,dx.$$

Así,

$$\left| \overline{\int_a^b} f(x)\,dx - \underline{\int_a^b} f(x)\,dx \right| \le 2\varepsilon(b-a);$$

Como esto es válido para todo $\varepsilon > 0$, las integrales inferior y superior son iguales, por lo que f es integrable. ■

5.3.2 Corolario *Supongamos que las funciones $g_k : [a, b] \to \mathbb{R}$ son integrables Riemann y que $\sum_{k=1}^{\infty} g_k$ converge uniformemente en $[a, b]$. Entonces podemos intercambiar el orden de la integración y de la suma:*

$$\int_a^b \left(\sum_{k=1}^{\infty} g_k(x) \right) dx = \sum_{k=1}^{\infty} \left(\int_a^b g_k(x)\,dx \right).$$

Demostración Sea $f_n = \sum_{k=1}^{n} g_k$; entonces $f_n \to f = \sum_{k=1}^{\infty} g_k$ (uniformemente) y, por **5.3.1**,

$$\int_a^b f_n(x)\,dx \to \int_a^b f(x)\,dx. \quad ■$$

5.3.3 Teorema *Sea $f_k :]a, b[\to \mathbb{R}$ una sucesión de funciones derivables en el intervalo abierto $]a, b[$ que convergen puntualmente a $f :]a, b[\to \mathbb{R}$. Supongamos que las derivadas f_k' son continuas y que convergen uniformemente a una función g. Entonces f es derivable y $f' = g$.*

Demostración Escribimos $f_k(x) = f_k(x_0) + \int_{x_0}^{x} f_k'(t)\,dt$, donde $a < x_0 < b$. Esto es posible por el teorema fundamental del cálculo. Si $k \to \infty$, obtenemos $f(x) = f(x_0) + \int_{x_0}^{x} g(t)\,dt$, por el teorema **5.3.1**. Como g es continua, por **5.1.4**, el teorema fundamental del cálculo muestra que el lado derecho es una función derivable de x con derivada $g(x)$. Por lo tanto, el miembro izquierdo es derivable, y entonces $f'(x) = g(x)$. ■

5.5.1 Teorema

i. *Si (M, d) y (N, ρ) son espacios métricos, entonces también lo es $C_b(A, N)$; es decir, $d(f, g)$ satisface*

a. $d(f, g) \ge 0$ *y* $d(f, g) = 0$ *sii* $f = g$.

b. $d(f, g) = d(g, f)$.

c. $d(f, g) \le d(f, h) + d(h, g)$.

ii. *Si N es un espacio normado, entonces también lo es $C_b(A, N)$; es decir, $\|\cdot\|$ satisface*

a. $\|f\| \geq 0$ *y* $\|f\| = 0$ *sii* $f = 0$.

b. $\|\alpha f\| = |\alpha| \, \|f\|$ *para* $\alpha \in \mathbb{R}, f \in C_b$.

c. $\|f + g\| \leq \|f\| + \|g\|$ *(desigualdad triangular)*.

Demostración

i. **a** y **b** son verificaciones rutinarias. Para demostrar **c**, escribimos

$$d(f, g) = \sup\{\rho(f(x), g(x)) \mid x \in M\}.$$

Para cada $x \in M$,

$$\rho(f(x), g(x)) \leq \rho(f(x), h(x)) + \rho(h(x), g(x))$$

y por la definición de supremo,

$$d(f, g) \leq \sup\{\rho(f(x), h(x)) + \rho(h(x), g(x)) \mid x \in M\}.$$

Entonces, $d(f, h) + d(h, g)$ es una cota superior del conjunto del miembro derecho, por lo que es mayor o igual que el supremo. Así, $d(f, g) \leq d(f, h) + d(h, g)$.

ii. **a** y **b** son claros. Para **c**,

$$\|f + g\| = \sup\{\|(f + g)(x)\| \mid x \in A\} \leq \sup\{\|f(x)\| + \|g(x)\| \mid x \in A\}$$

por la desigualdad triangular. Entonces, $\|f\| + \|g\|$ es una cota superior para el conjunto del miembro derecho, pues $\|f(x)\| \leq \|f\|$ y $\|g(x)\| \leq \|g\|$ para todo $x \in A$. Así, es mayor o igual que el supremo de este conjunto. Así, $\|f + g\| \leq \|f\| + \|g\|$, como se pedía. ■

5.5.2 Teorema *($f_k \to f$ (uniformemente en A)) $\Leftrightarrow$ ($f_k \to f$ en C_b).*

Demostración Esto no es más que reformular las definiciones. El estudiante debe proporcionar los detalles. ■

5.5.3 Teorema *Si N es un espacio métrico completo, entonces $C_b(A, N)$ es un espacio métrico completo. Si N es un espacio de Banach, también lo es $C_b(A, N)$.*

Demostración Sea $f_n \in C_b(A, N)$ una sucesión de Cauchy. Entonces f_n satisface el criterio de Cauchy (**5.2.1**) y, por lo tanto, converge uniformemente a una función f. Por **5.1.4**, f es continua, y si, por ejemplo, $\varepsilon = 1$ en la definición de convergencia, vemos que f_n acotada implica que f está acotada. Así, $f \in C_b(A, N)$. Por lo tanto, $C_b(A, N)$ es completo. La segunda afirmación es un caso particular de la primera. ■

5.6.2 Teorema de Arzela-Ascoli *Sean $A \subset M$ compacto y $\mathcal{B} \subset C_b(A, N)$. Entonces $\mathcal{B}$ es compacto si y sólo si $\mathcal{B}$ es cerrado, equicontinuo y puntualmente compacto.*

Demostración Para esta demostración, primero probaremos un lema.

Lema *Sea A compacto. Entonces, para cualquier $\delta > 0$ existe un conjunto finito $C_\delta = \{y_1, \ldots, y_k\}$ tal que cada $x \in A$ está a una distancia menor que δ de algún $y_i \in C_\delta$.*

Demostración La colección de bolas $\{D(x, \delta) \mid x \in A\}$ recubren a A, de modo que por compacidad, un número finito (digamos $D(y_1, \delta), \ldots, D(y_k, \delta)$) también lo recubren. Tómese $C\delta = \{y_1, \ldots, y_k\}$.

Pasemos ahora a la demostración del teorema. Sea $C = \cup\{C_{1/n} \mid n = 1, 2, 3, \ldots\}$. Como cada $C_{1/n}$ es finito, C es numerable; digamos que $C = \{x_1, x_2, \ldots\}$. Sea f_n nuestra sucesión en $\mathcal{B}$. Entonces $\{f_n\}$ está contenida en el conjunto puntualmente compacto $\mathcal{B}$ y el teorema de Bolzano-Weierstrass implica que existe una subsucesión de $f_n(x_1)$ que es convergente. Denotamos esta subsucesión

$$f_{11}(x_1), f_{12}(x_1), \ldots, f_{1n}(x_1), \ldots.$$

Análogamente, la sucesion $f_{1k}(x_2)$, $k = 1, 2, \ldots$, tiene una subsucesión

$$f_{21}(x_2), f_{22}(x_2), \ldots, f_{2n}(x_2), \ldots,$$

que es convergente. Continuando este proceso, la sucesión $f_{2k}(x_3)$, $k = 1, 2, \ldots$, tiene una subsucesión

$$f_{31}(x_3), f_{32}(x_3), \ldots, f_{3n}(x_3), \ldots,$$

que es convergente. Procedemos de esta forma y después hacemos $g_n = f_{nn}$, de modo que g_n es la n-ésima función que aparece en la n-ésima subsucesión.

En un diagrama, g_n se obtiene al elegir la diagonal

$$\begin{array}{llll}
f_{11}\ f_{12}\ f_{13}\cdots f_{1n}\cdots & \text{(primera subsucesión)} \\
f_{21}\ f_{22}\ f_{23}\cdots f_{2n}\cdots & \text{(segunda subsucesión)} \\
f_{31}\ f_{32}\ f_{33}\cdots f_{3n}\cdots & \text{(tercera subsucesión)} \\
\ \vdots\quad\ \vdots\quad\ \vdots\qquad\ \vdots & \\
f_{n1}\ f_{n2}\ f_{n3}\cdots f_{nn}\cdots & \text{(n-ésima subsucesión)}
\end{array}$$

Este truco se denomina "procedimiento diagonal" y es útil en muchas situaciones.

A partir de esta construcción, vemos que la sucesión g_n converge en cada punto de C; de hecho, g_n es subsucesión de cada sucesión f_{mk}, $k = 1, 2, \ldots$.

Ahora demostraremos que la sucesión g_n converge en cada punto de A y que la convergencia es uniforme, con lo que quedará demostrado el teorema. Para esto, sea $\varepsilon > 0$ y sea δ como en la definición de equicontinuidad. Sea $C_\delta = \{y_1, \ldots, y_k\}$ un subconjunto finito de C tal que todo punto de A está a una distancia menor que δ de algún punto en C_δ (véase el lema). Como todas las sucesiones

$$(g_n(y_1)), (g_n(y_2)), \ldots, (g_n(y_k))$$

convergen, existe un entero N tal que si $m, n \geq N$, entonces

$$\rho(g_n(y_i), g_n(y_i)) < \varepsilon \text{ para } i = 1, 2, \ldots, k.$$

Para cada $x \in A$, existe $y_j \in C_\delta$ tal que $\|x - y_j\| < \delta$. Por lo tanto, la hipótesis de equicontinuidad implica que

$$\rho(g_n(x), g_n(y_j)) < \varepsilon$$

para todo $n = 1, 2, \ldots$. Por lo tanto,

$$\begin{aligned}
\rho(g_n(x), g_m(x)) \leq\ & \rho(g_n(x), g_n(y_j)) + \rho(g_n(y_j), g_m(y_j)) \\
& + \rho(g_m(y_j), g_m(x)) \\
<\ & \varepsilon + \varepsilon + \varepsilon = 3\varepsilon,
\end{aligned}$$

siempre que $m, n \geq N$. Esto muestra que

$$\rho(g_n, g_m) \leq 3\varepsilon \quad \text{para} \quad m, n \geq N,$$

y por lo tanto, el criterio de Cauchy implica la convergencia uniforme de la sucesión g_n en A. El límite está en $\mathcal{B}$, pues $\mathcal{B}$ es cerrado. Esto demuestra que si $\mathcal{B}$ es equicontinuo y puntualmente compacto, entonces $\mathcal{B}$ es compacto. Dejaremos el recíproco al lector. (Véase también el ejercicio 68.) ■

5.6.3 Corolario *Sea $A \subset M$ compacto y $N = \mathbb{R}^m$. Supóngase que $\mathcal{B} \subset C(A, \mathbb{R}^m)$ es equicontinuo y puntualmente acotado. Entonces toda sucesión en $\mathcal{B}$ tiene una subsucesión uniformemente convergente.*

Demostración El conjunto de valores $\{f(x) \mid f \in \mathcal{B}\}$ para cada $x \in A$ fijo está acotado en $\mathbb{R}^n$, por lo que está dentro de un conjunto compacto (su clausura, por ejemplo). Así, la demostración anterior se aplica a una sucesión en $\mathcal{B}$.

Una alternativa es mostrar que $\text{cl}(\mathcal{B})$ es equicontinuo y puntualmente compacto, de modo que es compacto por **5.6.2**. ■

5.7.1 Principio de la aplicación contractiva *Sea M un espacio métrico completo y $\Phi : M \to M$ una transformación dada. Supóngase que existe una constante k, $0 \leq k < 1$ tal que $d(\Phi(x), \Phi(y)) \leq kd(x, y)$ para todo $x, y \in M$. Entonces existe un único punto fijo de Φ; es decir, un punto $x_* \in M$ tal que $(x_*) = x_*$. De hecho, si x_0 es cualquier punto en M y definimos $x_1 = \Phi(x_0), x_2 = \Phi(x_1), \ldots, x_{n+1} = \Phi(x_n), \ldots$, entonces*

$$\lim_{n \to \infty} x_n = x_*.$$

Demostración Primero mostramos la existencia de un punto fijo, y después su unicidad. Sean $x_0 \in M$ y $x_1, x_2, x_3, \ldots$ como en el teorema. Si $x_1 = x_0$, entonces $\Phi(x_0) = x_0$ y x_0 es un punto fijo. En caso contrario, $d(x_1, x_0)$ no es cero, por lo que comenzamos mostrando que los puntos $\{x_n\}$ forman una sucesión de Cauchy en M. Para mostrar esto, escribimos

$$d(x_2, x_1) = d(\Phi(x_1), \Phi(x_0)) \leq kd(x_1, x_0)$$
$$d(x_3, x_2) = d(\Phi(x_2), \Phi(x_1)) \leq kd(x_2, x_1) \leq k^2 d(x_1, x_0);$$

inductivamente, $d(x_n + 1, x_n) \leq k^n\, d(x_1, x_0)$. Además,

$$d(x_{n+p}, x_n) \leq d(x_{n+p}, x_{n+p-1}) + d(x_{n+p-1}, x_{n+p-2}) + \cdots + d(x_{n+1}, x_n)$$

por la desigualdad triangular, de modo que

$$d(x_{n+p}, x_n) \leq (k^{n+p-1} + k^{n+p-2} + \cdots + k^n) d(x_1, x_0).$$

Pero la serie geométrica $\sum_{i=0}^{\infty} k^i$ converge, pues $0 \leq k < 1$, por lo que satisface el criterio de Cauchy para series: dado $\varepsilon > 0$, existe N tal que

$$k^{n+p-1} + \cdots + k^n < \frac{\varepsilon}{d(x_1, x_0)}$$

si $n \geq N$ y p es arbitrario. Por lo tanto, $d(x_{n+p}, x_n) < \varepsilon$ si $n \geq N$ con p arbitrario, de modo que $\{x_n\}$ es una sucesión de Cauchy.

Por la completitud de M, $\lim_{n\to\infty} x_n$ existe en M. Llamamos a este límite x_*; es decir,

$$x_* = \lim_{n\to\infty} x_n.$$

Ahora mostramos que Φ es (uniformemente) continua. Dado $\varepsilon > 0$, sea $\delta = \varepsilon/k$. Entonces

$$d(x,y) < \delta \Rightarrow d(\Phi(x), \Phi(y)) \leq kd(x,y) < k\delta = \varepsilon.$$

Consideremos $x_{n+1} = \Phi(x_n)$; $x_{n+1} \to x_*$; por la continuidad de Φ, $\Phi(x_n) \to \Phi(x_*)$. Así, $x_* = \Phi(x_*)$, de modo que x_* es un punto fijo.

Por último, demostramos la unicidad del punto fijo x_*. Sea y_* otro punto fijo; es decir, $\Phi(y_*) = y_*$. Entonces

$$d(x_*, y_*) = d(\Phi(x_*), \Phi(y_*)) \leq kd(x_*, y_*); \text{ es decir, } (1-k)d(x_*, y_*) \leq 0.$$

Pero $k < 1$, por lo que $(1-k) > 0$, lo que implica $d(x_*, y_*) = 0$; es decir, $x_* = y_*$, y entonces el punto fijo es único. ∎

5.7.2 Teorema *Bajo las hipótesis del texto, existe $\delta > 0$ tal que la ecuación*

$$\frac{dx}{dt} = f(t,x), \quad x(t_0) = x_0 \tag{1}$$

tiene una única solución C^1 $x = \varphi(t)$, tal que $\varphi(t_0) = x_0$, para todo $t_0 - \delta < t < t_0 + \delta$; es decir, $\varphi'(t) = f(t, \varphi(t))$.

Demostración En primer lugar, elegimos una bola en torno a (t_0, x_0) tal que $|f(t,x)| \leq 1$ y donde se cumpla la condición de Lipschitz. Después elegimos δ tal que

i. $|t - t_0| \leq \delta$, $|x - x_0| \leq L\delta \Rightarrow (t,x)$ está en la bola elegida, y

ii. $K\delta < 1$ (donde K es la constante de Lipschitz).

Sea

$$\begin{aligned} M = \{\varphi : [t_0 - \delta, t_0 + \delta] \to \mathbb{R} \mid \varphi &\text{ es continua,} \\ \varphi(t_0) = x_0, \quad \text{y} \quad |\varphi(t) - x_0| &\leq L\delta\} \\ \subset C([t_0 - \delta, t_0 + \delta]). \end{aligned}$$

Obsérvese que M es cerrado en C y entonces M es un espacio métrico completo con métrica

$$d(\varphi_1, \varphi_2) = \sup_{t_0-\delta \leq t \leq t_0+\delta} |\varphi_1(t) - \varphi_2(t)|.$$

Sea $\Phi : M \to C$ dada por

$$\Phi(\varphi)(t) = x_0 + \int_{t_0}^{t} f(s, \varphi(s))\, ds.$$

Como hemos observado, la ecuación integral (2) en el texto es equivalente a (1) y ésta es equivalente a que φ sea un punto fijo de Φ. Mostraremos que Φ es una contracción.

En primer lugar, necesitamos demostrar que Φ aplica M en M. Es decir, si $\varphi \in M$, entonces $\psi(t) = x_0 + \int_{t_0}^{t} f(s, \varphi(s))\, ds \in M$. La función ψ es continua, pues satisface una condición de Lipschitz con constante L: Si $s < t$, entonces

$$|\psi(t) - \psi(s)| = \left|\int_s^t f(x, \varphi(x))\, dx\right| \leq \int_s^t |f(x, \varphi(x))|\, dx \leq L|t - s|.$$

Obsérvese que $\psi(t_0) = x_0$ y

$$|\psi(t) - x_0| = \left|\int_{t_0}^t f(s, \varphi(s))\, ds\right| \leq L|t - t_0| \leq L\delta,$$

por lo que de hecho $\psi \in M$. A continuación buscamos k, $0 \leq k < 1$, tal que $d(\Phi(\varphi_1), \Phi(\varphi_2)) \leq kd\,(\varphi_1, \varphi_2)$. Observamos que

$$d(\Phi(\varphi_1), \Phi(\varphi_2)) = \sup_t \left|\int_{t_0}^t [f(s, \varphi_1(s)) - f(s, \varphi_2(s))]\, ds\right|$$
$$\leq \delta K \sup_s |\varphi_1(s) - \varphi_2(s)| = \delta K d(\varphi_1, \varphi_2)$$

y δ se eligió de modo que $\delta K < 1$. Así, tomamos $k = \delta K$.

En conclusión, como Φ es una contracción, tiene un único punto fijo, lo que muestra que nuestra ecuación tiene una única solución. ■

5.7.3 Proposición *Si* $\lambda M\,|b - a| < 1$, *entonces la ecuación de Fredholm (3) tiene una única solución.*

Demostración Dimos toda la demostración en el texto, excepto el hecho de que $\Phi(f)$ es continua. Para demostrar esto, obsérvese que, como $[a, b] \times [a, b]$ es compacto, K es uniformemente continua. Por lo tanto, dado $\varepsilon > 0$, existe $\delta > 0$ tal que $((x - y)^2 + (t - s)^2)^{1/2} < \delta$ implica $|K(x, t) - K(y, s)| < \varepsilon$. En particular, $|x - y| < \delta$ implica $|K(x, t) - K(y, t)| < \varepsilon$, para todo $t \in [a, b]$. Así,

$$\int_a^b |K(x, t) - K(y, t)|\,|f(t)|\, dt < \varepsilon \int_a^b |f(t)|\, dt < \varepsilon\|f\|(b - a)$$

y entonces

$$|(\Phi(f))(x) - (\Phi(f))(y)| < \varepsilon\lambda(b - a)\|f\|$$

lo que implica que $\Phi(f)$ es continua. ■

5.7.4 Proposición *Supóngase que K es continua en $[a, b] \times [a, b]$; entonces, la ecuación integral de Volterra tiene una única solución $f(x)$ para* **cualquier** λ.

Demostración Primero preparamos un útil lema.

Lema *Sean M un espacio métrico completo y $\Phi : M \to M$ tal que para algún entero positivo N, Φ^N, definida como $\Phi^N = \Phi \circ \Phi \circ \cdots \circ \Phi$ (N veces) es una contracción. Entonces Φ tiene un único punto fijo.*

Demostración Sea x_* el único punto fijo de Φ^N. Entonces $\Phi^N(x_*) = x_*$, de modo que $\Phi^{N+1}(x_*) = \Phi(x_*)$ o $\Phi^N(\Phi(x_*)) = \Phi(x_*)$. Así, $\Phi(x_*)$ también es un punto fijo de Φ^N y, por unicidad, $\Phi^N(x_*) = x_*$. Así, Φ tiene un punto fijo. Cualquier punto fijo de Φ también es un punto fijo de Φ^N, y por lo tanto, es único. ∇

Demostración de 5.7.4 Sea $\Phi : C([a, b], \mathbb{R}) \to C([a, b], \mathbb{R})$ definida por

$$\Phi(f)(x) = \lambda \int_a^x K(x, y) f(y)\, dy + \varphi(x).$$

Como en **5.7.3**, $\Phi(f)$ es continua. Obsérvese que

$$\begin{aligned} d(\Phi(f_1), \Phi(f_2)) &= \sup |\Phi(f_1)(x) - \Phi(f_2)(x)| \leq \lambda M \sup |x - a| d(f_1, f_2) \\ &= \lambda M |b - a| d(f_1, f_2). \end{aligned}$$

Además,

$$\begin{aligned} d(\Phi^2(f_1), \Phi^2(f_2)) &= d(\Phi(\Phi(f_1)), \Phi(\Phi(f_2))) \\ &\leq \sup_x \left| \lambda \int_a^x K(x, y) \lambda\, M|y - a| d(f_1, f_2)\, dy \right| \\ &\leq \lambda^2 M^2 \left(\sup_x \frac{|x - a|^2}{2} \right) d(f_1, f_2). \end{aligned}$$

Inductivamente,

$$d(\Phi^n(f_1), \Phi^n(f_2)) \leq \frac{\lambda^n M^n |b - a|^n}{n!} d(f_1, f_2).$$

Por el criterio del cociente, $\sum_{n=0}^{\infty} \lambda^n M^n |b - a|^n / n!$ converge y, en consecuencia, existe un entero N tal que $\lambda^N M^N |b - a|^N / N! < 1$, lo que implica que Φ^N es una contracción. Por el lema, podemos concluir que Φ tiene un único punto fijo. ■

5.8.1 Teorema *Sean $f\colon [0,1] \to \mathbb{R}$ continua y $\varepsilon > 0$. Entonces existe un polinomio $p(x)$ tal que $\|p - f\| < \varepsilon$. De hecho, la sucesión de polinomios de Bernstein*

$$p_n(x) = \sum_{k=0}^{n} \binom{n}{k} f(k/n) x^k (1-x)^{n-k}$$

converge uniformemente a f si $n \to \infty$, donde

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

denota el coeficiente binómico.

Demostración El teorema del binomio afirma que

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}. \tag{1}$$

Al derivar la ecuación (1) con respecto a x y multiplicarla por x se obtiene la identidad

$$nx(x+y)^{n-1} = \sum_{k=0}^{n} k \binom{n}{k} x^k y^{n-k}. \tag{2}$$

Análogamente, al derivar dos veces,

$$n(n-1)x^2(x+y)^{n-2} = \sum_{k=0}^{n} k(k-1) \binom{n}{k} x^k y^{n-k}. \tag{3}$$

Sea (por simplicidad de notación)

$$r_k(x) = \binom{n}{k} x^k (1-x)^{n-k}.$$

Así, las ecuaciones (1), (2) y (3) se leen (con $y = 1 - x$) como

$$\sum_{k=0}^{n} r_k(x) = 1, \quad \sum_{k=0}^{n} k r_k(x) = nx, \quad \sum_{k=0}^{n} k(k-1) r_k(x) = n(n-1)x^2.$$

Esto implica que tenemos la identidad

$$\begin{aligned}\sum_{k=0}^{n} (k - nx)^2 r_k(x) &= n^2x^2 \sum_{k=0}^{n} r_k(x) - 2nx \sum_{k=0}^{n} k r_k(x) + \sum_{k=0}^{n} k^2 r_k(x) \\ &= n^2x^2 - 2nx \cdot nx + [nx + n(n-1)x^2] \\ &= nx(1-x). \end{aligned} \tag{4}$$

Ahora elegimos M tal que $|f(x)| \leq M$ en $[0, 1]$. Como f es uniformemente continua, dado $\varepsilon > 0$ existe $\delta > 0$ tal que $|x - y| < \delta$ implica $|f(x) - f(y)| < \varepsilon$.

Queremos estimar la expresión

$$|f(x) - p_n(x)| = \left| f(x) - \sum_{k=0}^{n} f(k/n) r_k(x) \right| = \left| \sum_{k=0}^{n} (f(x) - f(k/n))\, r_k(x) \right|.$$

Para esto, dividimos esta suma en dos partes: una con los puntos tales que $|k - nx| < \delta n$ y otra con los puntos tales que $|k - nx| \geq \delta n$. Si $|k - nx| < \delta n$, entonces $|x - (k/n)| < \delta$, de modo que $|f(x) - f(k/n)| < \varepsilon$ y en consecuencia, recordando que $r_k(x) \geq 0$, estos términos dan una suma $\leq \varepsilon \sum_{k=0}^{n} r_k(x) = \varepsilon$. Los términos del segundo tipo tienen una suma

$$\leq 2M \sum_{|k-nx| \geq \delta n} r_k(x) \leq \frac{2M}{n^2\delta^2} \sum_{k=0}^{n} (k - nx)^2 r_k(x),$$

pues $|(k - nx)/n\delta| \geq 1$ para estos términos. Por (4), esta suma está acotada por

$$\frac{2Mx(1 - x)}{n\delta^2} \leq \frac{M}{2\delta^2 n},$$

pues $x(1 - x) \leq 1/4$ (¿por qué?). Así, hemos demostrado que para cada $\varepsilon > 0$ existe $\delta > 0$ tal que

$$|f(x) - p_n(x)| < \varepsilon + \frac{M}{2\delta^2 n}.$$

Así, si $n \geq M/2\delta^2\varepsilon$, entonces $M/2\delta^2 n < \varepsilon$, de modo que $|f(x) - p_n(x)| < 2\varepsilon$; así, $p_n \to f$ uniformemente. ■

5.8.2 Teorema de Stone-Weierstrass *Sean M un espacio métrico, $A \subset M$ un conjunto compacto y $B \subset C(A, \mathbb{R})$, con las siguientes condiciones:*

i. *B es un álgebra; si $f, g \in B$ y $\alpha \in \mathbb{R}$, entonces $f + g \in B$, $f \cdot g \in B$ y $\alpha f \in B$.*

ii. *La función constante $x \mapsto 1$ está en B.*

iii. *B separa puntos; es decir, para $x, y \in A$, $x \neq y$, existe $f \in B$ tal que $f(x) \neq f(y)$.*

Entonces B es denso en $C(A, \mathbb{R})$; es decir, $\mathrm{cl}(B) = C(A, \mathbb{R})$.

Demostración Utilizaremos la siguiente notación:

$$(f \vee g)(x) = \text{máx}(f(x), g(x)) \quad \text{y} \quad (f \wedge g)(x) = \text{mín}(f(x), g(x)).$$

(Véase la figura 5.P-1.) Sea $\overline{\mathcal{B}}$ la clausura de $\mathcal{B}$. Entonces, por la continuidad de la suma y la multiplicación, vemos que $\overline{\mathcal{B}}$ también satisface **i**. Es claro que satisface **ii** y **iii**. Así, queremos demostrar que $\overline{\mathcal{B}} = C(A, \mathbb{R})$.

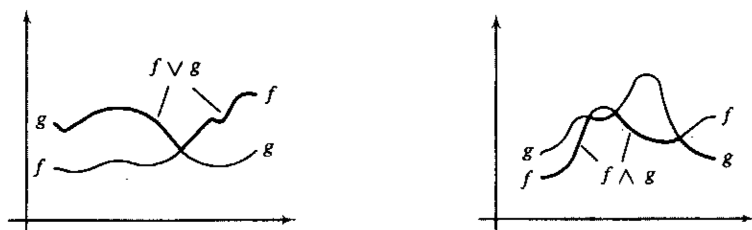


FIGURA 5.P-1 El significado gráfico de $f \vee g$ y $f \wedge g$

Por el teorema anterior y el ejemplo **5.8.4**, podemos encontrar una sucesión de polinomios $p_n(t)$ tales que

$$\big||t| - p_n(t)\big| < 1/n \quad \text{para} \ -n \leq t \leq n.$$

Así,

$$\big||f(x)| - p_n(f(x))\big| < 1/n \quad \text{si} \ -n \leq f(x) \leq n.$$

Esto demuestra que para $f \in \overline{\mathcal{B}}$, $|f| \in \overline{\mathcal{B}}$, pues $p_n \circ f \in \overline{\mathcal{B}}$, ya que $\overline{\mathcal{B}}$ es un álgebra.

Ahora tenemos las identidades

$$f \vee g = \frac{f+g}{2} + \frac{|f-g|}{2} \quad \text{y} \quad f \wedge g = \frac{f+g}{2} - \frac{|f-g|}{2}$$

(un ejercicio para el lector). Así, si $f, g \in \overline{\mathcal{B}}$, entonces $f \vee g$ y $f \wedge g$ también están en $\overline{\mathcal{B}}$.

Sean $h \in C(A, \mathbb{R})$ y $x_1, x_2 \in A$, con $x_1 \neq x_2$. Elegimos $g \in \mathcal{B}$ tal que $g(x_1) \neq g(x_2)$ (lo que es posible por la hipótesis **iii**), y sea $fx_1x_2(x) = \alpha g(x) + \beta$, donde

$$\alpha = \frac{[h(x_1) - h(x_2)]}{[g(x_1) - g(x_2)]} \quad \text{y} \quad \beta = \frac{[g(x_1)h(x_2) - h(x_1)g(x_2)]}{[g(x_1) - g(x_2)]}.$$

Elegimos estos números α y β de modo que $f_{x_1x_2}(x_1) = h(x_1)$ y $f_{x_1x_2}(x_2) = h(x_2)$.

Sean $\varepsilon > 0$ y $x \in A$. Para $y \in A$, existe una vecindad $U(y)$ de y tal que $f_{yx}(z) > h(z) - \varepsilon$ si $z \in U(y)$; esto se debe simplemente a la continuidad de h. Sea $U(y_1), \ldots, U(y_r)$ un subrecubrimiento finito de A, que existe gracias a que A es compacto. Tomemos $f_x = f_{y_1x} \vee \cdots \vee f_{y_rx}$. Así, como antes, $f_x \in \overline{B}$ y $f_x(z) > h(z) - \varepsilon$ para todo $z \in A$. Además, $f_x(x) = h(x)$. Así, existe una vecindad $V(x)$ tal que $f_x(y) < h(y) + \varepsilon$ si $y \in V(x)$. Sea $V(x_1), \ldots, V(x_k)$ un recubrimiento de A y tomemos $f = f_{x_1} \wedge \cdots \wedge f_{x_k}$. De nuevo, $f \in \overline{B}$. Ahora, $f(z) > h(z) - \varepsilon$ para todo $z \in A$, pues $f_{x_j}(u) > h(u) - \varepsilon$ para todo $u \in A$. Dado $y \in A$, $y \in V(x_j)$ para algún x_j, por lo que $f(y) \le f_{x_j}(y) < h(y) + \varepsilon$. Entonces $|f(z) - h(z)| < \varepsilon$ y $h \in \overline{B}$. Así, $\overline{B} = C(A, \mathbb{R})$. ■

Las demostraciones de §**5.9** usan la ***fórmula de sumación parcial de Abel***:

Considérense dos sucesiones $a_1, a_2, \ldots$ y $b_1, b_2, \ldots$ de números reales y sea $s_n = a_1 + \cdots + a_n$. Entonces

$$\sum_{k=1}^{n} a_k b_k = s_n b_{n+1} - \sum_{k=1}^{n} s_k(b_{k+1} - b_k) = s_n b_1 + \sum_{k=1}^{n}(s_n - s_k)(b_{k+1} - b_k).$$

Demostración Obsérvese que $a_n = s_n - s_{n-1}$. Entonces

$$\sum_{k=1}^{n} a_k b_k = \sum_{k=1}^{n}(s_k - s_{k-1})b_k = \sum_{k=1}^{n} s_k b_k - \sum_{k=1}^{n} s_{k-1} b_k,$$

donde $s_0 = 0$, de donde

$$\sum_{k=1}^{n} s_{k-1} b_k = \sum_{k=1}^{n} s_k b_{k+1} - s_n b_{n+1},$$

por lo que obtenemos el primer resultado. La segunda igualdad se obtiene al poner

$$b_{n+1} = \sum_{k=1}^{n}(b_{k+1} - b_k) + b_1$$

en la primera igualdad. ■

5.9.1 Criterio de Abel *Sean $A \subset \mathbb{R}^m$ y $\varphi_n : A \to \mathbb{R}$ una sucesión decreciente de funciones; es decir, $\varphi_{n+1}(x) \le \varphi_n(x)$ para cada $x \in A$. Supóngase que existe una constante M tal que $|\varphi_n(x)| \le M$ para todo $x \in A$ y todo n. Si $\sum_{n=1}^{\infty} f_n(x)$ converge uniformemente en A, entonces también lo hace $\sum_{n=1}^{\infty} \varphi_n(x) f_n(x)$.*

Demostración Sean

$$s_n(x) = \sum_{k=1}^{n} f_k(x) \quad \text{y} \quad r_n(x) = \sum_{k=1}^{n} \varphi_k(x)\, f_k(x).$$

Por la segunda igualdad del lema, tenemos que

$$r_n(x) - r_m(x) = (s_n(x) - s_m(x))\varphi_{m+1}(x) + \sum_{k=m+1}^{n} (s_n(x) - s_k(x))(\varphi_{k+1}(x) - \varphi_k(x))$$

para $n \geq m$, de modo que

$$\begin{aligned} |r_n(x) - r_m(x)| \leq\ & |s_n(x) - s_m(x)|\ |\varphi_{m+1}(x)| \\ & + \sum_{k=m+1}^{n} |s_n(x) - s_k(x)|\ |\varphi_{k+1}(x) - \varphi_k(x)|, \end{aligned}$$

pues

$$\varphi_{k+1} \leq \varphi_k \quad \text{y} \quad |\varphi_{k+1} - \varphi_k| = \varphi_k - \varphi_{k+1}.$$

Dado $\varepsilon > 0$, sea N tal que $n, m \geq N$ implica $|s_n(x) - s_m(x)| < \varepsilon/3M$ para todo $x \in A$. Entonces, para todo $x \in A$,

$$\begin{aligned} |r_n(x) - r_m(x)| &< \frac{\varepsilon}{3} + \left(\frac{\varepsilon}{3M}\right) \sum_{k=m+1}^{n} [\varphi_k(x) - \varphi_{k+1}(x)] \\ &= \frac{\varepsilon}{3} + \left(\frac{\varepsilon}{3M}\right) [\varphi_{m+1}(x) - \varphi_{n+1}(x)] \\ &\leq \frac{\varepsilon}{3} + \left(\frac{\varepsilon}{3M}\right) [|\varphi_{m+1}(x)| + |\varphi_{n+1}(x)|] \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

Por lo tanto, el criterio de Cauchy implica que $r_n(x)$ converge uniformemente. ■

5.9.2 Criterio de Dirichlet *Sea $s_n(x) = \sum_{i=1}^{n} f_i(x)$ para una sucesión $f_n : A \subset \mathbb{R}^m \to \mathbb{R}$. Supóngase que existe M constante tal que $|s_n(x)| \leq M$ para todo $x \in A$ y todo n. Sea $g_n : A \subset \mathbb{R}^m \to \mathbb{R}$ tal que $g_n \to 0$ uniformemente, $g_n \geq 0$ y $g_{n+1}(x) \leq g_n(x)$. Entonces $\sum_{n=1}^{\infty} f_n(x) g_n(x)$ converge uniformemente en A.*

Demostración Usaremos la misma notación que en la demostración del criterio de Abel, $\varphi_n = g_n$. Sin embargo, para calcular $r_n - r_m$, ahora usamos la primera igualdad del lema, a saber,

$$r_n(x) - r_m(x) = s_n(x)\varphi_{n+1}(x) - s_m(x)\varphi_{m+1}(x) - \sum_{k=m+1}^{n} s_k(x)(\varphi_{k+1}(x) - \varphi_k(x)),$$

de modo que, como $\varphi_k \leq 0$ y $\varphi_{k+1} \leq \varphi_k$,

$$\begin{aligned} |r_n(x) - r_m(x)| &\leq M(\varphi_{n+1}(x) - \varphi_{m+1}(x)) + M \sum_{k=m+1}^{n} (\varphi_k(x) - \varphi_{k+1}(x)) \\ &= M(\varphi_{n+1}(x) + \varphi_{m+1}(x) + \varphi_{m+1}(x) - \varphi_{n+1}(x)) = 2M\varphi_{m+1}(x). \end{aligned}$$

Dado $\varepsilon > 0$, elegimos N tal que $m > N$ implica $\varphi_m(x) < \varepsilon/2M$ para todo x. Entonces, $m, n \geq N$ implica $|r_n(x) - r_m(x)| < \varepsilon$, lo que demuestra la afirmación. ■

5.10.2 Teorema *$\sum_{k=0}^{\infty} a_k x^k (x - x_0)^k$ converge absolutamente para $|x - x_0| < R$, converge uniformemente para $|x - x_0| \leq R'$, donde $R' < R$ y diverge si $|x - x_0| > R$ (el teorema no proporciona información si $|x - x_0| = R$).*

Demostración Recuérdese de **5.10.1** que $R^{-1} = \text{lím sup} \sqrt[k]{|a_k|}$ cuando $k \to \infty$. Sea $R' < R$. Elegimos R'' tal que $R' < R'' < R$. Para n suficientemente grande, la definición del límite superior implica

$$\sqrt[n]{|a_n|} \leq \frac{1}{R''}, \quad \text{es decir,} \quad |a_n| \leq \left(\frac{1}{R''}\right)^n.$$

Por lo tanto, si $|x - x_0| \leq R'$, tenemos que

$$|a_n(x - x_0)^n| \leq \left(\frac{R'}{R''}\right).$$

Como $R'/R'' < 1$, tenemos la convergencia uniforme y absoluta de la serie en el disco $|x - x_0| \leq R'$, por el criterio M de Weierstrass.

Por otro lado, supongamos que la serie $\sum a_n(x - x_0)^n$ converge. Entonces $a_n(x - x_0)^n \to 0$, de modo que $|a_n(x - x_0)^n| \leq 1$ para n grande. Así, $\sqrt[n]{|a_n|} \leq |x - x_0|^{-1}$ para n grande. Por lo tanto, $R^{-1} = \text{lím sup} \sqrt[n]{|a_n|} \leq |x - x_0|^{-1}$; es decir, $|x - x_0| \leq R$. ■

5.10.3 Corolario *La suma de una serie de potencias es una función C^∞ dentro de su círculo de convergencia. Se puede derivar término a término; la serie derivada tiene el mismo radio de convergencia.*

Demostración La serie que se obtiene al derivar término a término la serie dada es $\sum ka_k(x-x_0)^{k-1}$, cuyo radio de convergencia es R', donde

$$\frac{1}{R'} = \text{lím sup} \sqrt[k]{k|a_k|}.$$

Pero $\sqrt[k]{k} \to 1$ (¿por qué?), de modo que

$$\frac{1}{R'} = \text{lím sup} \sqrt[k]{k|a_k|} = \frac{1}{R}, \quad \text{es decir,} \quad R' = R.$$

Así, la serie derivada converge uniformemente dentro de cualquier círculo menor, y por lo tanto es la derivada de la suma de la serie original (véase el corolario **5.3.4**). Por inducción, vemos que la serie original es derivable cualquier número de veces. ■

5.10.6 Teorema (Abel) *Si* $\sum_{k=0}^{\infty} a_k = A$, *entonces* $\sum_{k=0}^{\infty} a_k x^k$ *converge para* $|x| < 1$ *y* $\text{lím}_{x\to 1^-} \sum_{k=0}^{\infty} a_k x^k = A$.

Demostración Si cambiamos a_0, podemos suponer que $A = 0$. Como a_k está acotada (de hecho, $a_k \to 0$), la serie $\sum_{k=0}^{\infty} a_k x^k$ converge para $|x| < 1$, por el teorema **5.10.2**.

Escribamos $S_n = \sum_{k=0}^{n} a_k$. Como S_n está acotada cuando $n \to \infty$, la serie $\sum_{k=0}^{\infty} S_k x^k$ converge igualmente para $|x| < 1$. Como $A = 0$, $S_n \to 0$ cuando $n \to \infty$. Escribamos $f(x) = \sum_{k=0}^{\infty} a_k x^k$ para $|x| < 1$. Entonces

$$f(x) = S_0 + \sum_{k=1}^{\infty} (S_k - S_{k-1})x^k = (1-x)\sum_{k=1}^{\infty} S_k x^k.$$

Como $S_n \to 0$, dado $\varepsilon > 0$ podemos determinar n_0 tal que $|S_n| \le \varepsilon$ para $n > n_0$. Entonces

$$\begin{aligned}
|f(x)| &\le (1-x)\left|\sum_{k=1}^{n_0} S_k x^k\right| + (1-x)\sum_{k=n_0+1}^{\infty} \varepsilon x^k \\
&\le (1-x)\left|\sum_{k=1}^{n_0} S_k x^k\right| + (1-x)\cdot \varepsilon x^{n_0+1}(1-x)^{-1} \\
&\le (1-x)\left|\sum_{k=1}^{n_0} S_k x^k\right| + \varepsilon.
\end{aligned}$$

Por lo tanto, $\text{lím sup}_{x\to 1^-} |f(x)| \le \varepsilon$. Como $\varepsilon > 0$ era arbitrario, $\text{lím sup}_{x\to 1^-} f(x) = 0$. ■

5.10.7 Teorema *$\sum_{k=0}^{\infty} a_k = A$ (C, 1) implica $\sum_{k=0}^{\infty} a_k = A$ (Abel).*

Demostración Como antes, podemos suponer que $A = 0$. Escribimos $S_n = \sum_{k=0}^{n} a_k$, $T_n = \sum_{k=0}^{n} S_k$. Por hipótesis, $T_n = o(n)$; aquí usamos la notación "o minúscula", que indica que para cada $\varepsilon > 0$, $|o(n)| \leq \varepsilon n$ para n suficientemente grande. Por lo tanto, $S_n = T_n - T_{n-1} = o(n)$ y $a_n = S_n - S_{n-1} = o(n)$. Mediante una comparación con un múltiplo de la serie convergente $\sum kx^k$, las tres series $\sum a_k x^k$, $\sum S_k x^k$, y $\sum T_k x^k$ convergen si $|x| < 1$. Además,

$$f(x) = \sum a_k x^k = (1-x)\sum S_k x^k = (1-x)^2 \sum T_k x^k.$$

Ahora, como $T_n = o(n)$, dado $\varepsilon > 0$ podemos elegir n_0 tal que $n \geq n_0$ implique $|T_n| \leq \varepsilon n$. En consecuencia,

$$\begin{aligned} |f(x)| &\leq (1-x)^2 \left| \sum_{k \leq n_0} T_k x^k \right| + (1-x)^2 \left| \sum_{k > n_0} \varepsilon k x^k \right| \\ &\leq (1-x)^2 \left| \sum_{k \leq n_0} T_k x^k \right| + (1-x)^2 \cdot \varepsilon x (1-x)^{-2} \end{aligned}$$

y vemos que $\lim \sup_{x \to 1^-} |f(x)| \leq \varepsilon$. Así, como en el teorema anterior, $\lim \sup_{x \to 1^-} f(x) = 0$. ■

5.10.8 Teorema *Si $\sum a_n = A$ (C, 1) y $a_n = O(1/n)$ (es decir, si $|a_n| \leq C/n$ para C constante), entonces $\sum a_n$ converge (a A) en el sentido usual.*

Demostración Como en las demostraciones anteriores, podemos suponer que $A = 0$. Escribimos $S_n = \sum_{k=1}^{n} a_k$, $T_n = \sum_{k=1}^{n} S_k$. Entonces, la primera hipótesis se escribe $T_n = o(n)$, como en la demostración anterior.

Queremos demostrar que $S_n \to 0$. En caso contrario, para algún $\delta > 0$, $|S_n| \geq \delta$ para un conjunto infinito de índices n. Podemos suponer (invirtiendo todos los signos en caso necesario) que $S_n \geq \delta$ para infinitos valores de n. Pero si $S_n \geq \delta$ y $r > n$, tenemos que

$$S_r = S_n + a_{n+1} + a_{n+2} + \cdots + a_r \geq \delta - C\left(\frac{1}{n+1} + \cdots + \frac{1}{R}\right) \geq \delta - C \log \frac{r}{n}.$$

Esto será mayor o igual que $\delta/2$ si $C \log(r/n) \leq \delta/2$; es decir, $r/n \leq e^{\delta/2C} = \lambda$.

(Obsérvese que $\lambda > 1$.) Por lo tanto, tenemos

$$([\lambda\ n] - n)\frac{\delta}{2} \leq \sum_{r=n+1}^{[\lambda n]} S_r = T_{[\lambda n]} - T_n.$$

(Aquí $[x]$ indica el mayor entero menor o igual que x.) El miembro derecho de esta desigualdad es $o(n)$, pero el izquierdo es del orden $(\lambda - 1)\delta n/2$, una contradicción. Por lo tanto, S_n debe tender a cero. ■

Ejemplos resueltos del capítulo 5

Ejemplo 5.1

a. *Si $f_k \to f$ (puntualmente) y $g_k \to g$ (puntualmente), muéstrese que $f_k + g_k \to f + g$ (puntualmente) para cualesquiera funciones $f, g : A \subset \mathbb{R}^n \to \mathbb{R}^m$.*

b. *Respóndase la misma pregunta para la convergencia uniforme.*

Solución

a. Para $x \in A$, debemos mostrar que $(f_k + g_k)(x) \to (f + g)(x)$. Dado $\varepsilon > 0$, elegimos N_1 tal que $k \geq N_1$ implique $\|f_k(x) - f(x)\| < \varepsilon/2$ y N_2 tal que $k \geq N_2$ implique $\|g_k(x) - g(x)\| < \varepsilon/2$. Entonces, sea $N = \text{máx}(N_1, N_2)$, de modo que $k \geq N$ implique (por la desigualdad triangular)

$$\|(f_k + g_k)(x) - (f + g)(x)\| \leq \|f_k(x) - f(x)\| + \|g_k(x) - g(x)\| < \varepsilon.$$

b. Se repite el argumento en **a**, donde cada afirmación es válida para todo $x \in A$. ♦

Ejemplo 5.2 *Demuéstrese que una sucesión $f_k : A \to \mathbb{R}^n$ converge puntualmente (uniformemente) sii sus componentes convergen puntualmente (uniformemente).*

Solución La parte del ejemplo correspondiente a la convergencia puntual es consecuencia del hecho de que una sucesión en $\mathbb{R}^m$ converge sii sus componentes lo hacen (véase el capítulo 2). Sin embargo, escribiremos de nuevo el argumento para ver la validez de la convergencia uniforme.

Sea $x = (x^1, \ldots, x^m) \in \mathbb{R}^m$. Entonces $|x^i| \leq \|x\| \leq \sum_{i=1}^m |x^i|$. De hecho, la primera desigualdad es evidente y la segunda es consecuencia de la desigualdad triangular si escribimos $x = (x^1, 0, \ldots, 0) + (0, x^2, 0, \ldots, 0) + \cdots + (0, 0, \ldots, x^m)$.

Si aplicamos esto a $f_k = (f_k^1, \ldots, f_k^m)$, tenemos

$$|f_k^i(x) - f^i(x)| \leq \|f_k(x) - f(x)\| \leq \sum_{i=1}^m |f_k^i(x) - f^i(x)|.$$

Por lo tanto, si f_k converge puntualmente, entonces f_k^i converge puntualmente.

Recíprocamente, supóngase que $f_k^i(x)$ converge para cada i y x. Elíjase N_i tal que $k, l \geq N_i$ implique $|f_k^i(x) - f_l^i(x)| < \varepsilon/m$. Entonces, si $N = \text{máx}(N_1, \ldots, N_m)$, $k, l \geq N$, implica $\|f_k(x) - f_l(x)\| < \varepsilon/m + \cdots + \varepsilon/m = \varepsilon$, de modo que $f_k(x)$ converge.

Para la convergencia uniforme, repetimos el argumento, donde cada afirmación es válida para todo $x \in A$. ♦

Ejemplo 5.3 *Encuéntrese un ejemplo de sucesión f_k que converja uniformemente a cero en $[0, \infty[$, donde cada $\int_0^\infty f_k(x)\, dx$ exista, pero $\int_0^\infty f_k(x)\, dx = +\infty$. ¿Contradice esto el teorema* **5.3.1**?

Solución Sea

$$f_k(x) = \begin{cases} 1/k, & \text{si } 0 \leq x \leq k^2, \\ 0, & \text{si } x > k^2. \end{cases}$$

Entonces $f_k \to 0$ uniformemente, pues $|f_k(x)| \leq 1/k$ para todo x. Sin embargo,

$$\int_0^\infty f_k(x)\, dx = k^2/k = k \to \infty.$$

Esto no contradice el teorema **5.3.1** pues el teorema se refiere a intervalos finitos. ♦

Ejemplo 5.4 (Teorema de Dini) *Sea $A \subset \mathbb{R}^n$ compacto y f_k una sucesión de funciones continuas $f_k : A \to \mathbb{R}$ tales que* (**a**) *$f_k(x) \geq 0$ para $x \in A$;* (**b**) *$f_k \to 0$ puntualmente; y* (**c**) *$f_k(x) \leq f_l(x)$ si $k \geq l$. Demuéstrese que $f_k \to 0$ uniformemente.*

Solución Este ejemplo requiere un poco de cuidado, pues tratamos de deducir la convergencia uniforme a partir de la convergencia puntual más alguna hipótesis adicional, y sabemos que el resultado no será cierto sin esa hipótesis adicional (estúdiese la figura 5.1-1, donde todas estas hipótesis son válidas, excepto $f_k(0) \to 0$ cuando $k \to \infty$).

Dado $\varepsilon > 0$, queremos encontrar N tal que $|f_k(x)| < \varepsilon$ para todo $k \geq N$ y $x \in A$. Para cada $x \in A$, encontramos N_x tal que $|f_k(x)| < \varepsilon/2$ si $k \geq N_x$. Escribimos N_x para enfatizar que este número depende de x. Aquí hemos utilizado la hipótesis (b). Por la continuidad de $f_k(x)$, existe una vecindad $U_{x,k}$ de x tal que $|f_k(y) - f_k(x)| < \varepsilon/2$ para $y \in U_{x,k}$. Las vecindades U_{x,N_x} forman un recubrimiento de A; por la compacidad, existe un subrecubrimiento finito, digamos, centrado en $x_1, \ldots, x_M$. Sea $N = \text{máx}(N_{x_1}, \ldots, N_{x_M})$, y sean $x \in A$ y $k \geq N$. Entonces $x \in U_{x_l,N_l}$ para algún l, de modo que $|f_{N_l}(x) - f_{N_l}(x_l)| < \varepsilon/2$. Así, usando (a) y (c),

$$\begin{aligned} 0 \leq f_k(x) &\leq f_N(x) \\ &\leq f_{N_l}(x) = f_{N_l}(x_l) + [f_{N_l}(x) - f_{N_l}(x_l)] \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

Por lo tanto, $|f_k(x)| < \varepsilon$ para todo $k \geq N$ *y todo* $x \in A$, y así tenemos la convergencia uniforme. ♦

Ejemplo 5.5 Considérese la serie alternante $\sum_{n=1}^{\infty} (-1)^n/n$, la cual converge (véase el teorema **5.9.2**). No podemos reordenar los términos de la serie, pues puede ocurrir que diverja. ¡De hecho, la serie $\sum(-1)^n/n$ se puede reordenar para obtener cualquier suma! Este hecho fue descubierto por Riemann (véase el ejercicio 17). Para poder reordenar una serie, es suficiente tener convergencia absoluta. En primer lugar, definamos una reordenación. Sea $\sum_{i=1}^{\infty} a_i$ una serie. Una ***reordenación*** es la serie $\sum_{n=1}^{\infty} a_{\sigma(i)}$, donde σ es una permutación de $\{1, 2, 3, \ldots\}$, o, más precisamente, una biyección $\sigma: \{1, 2, 3, \ldots\} \to \{1, 2, 3, \ldots\}$.

Demuéstrese el siguiente teorema.

Teorema *Sea $g_k \in \mathbb{R}^m$ y supóngase que $\sum_{k=1}^{\infty} g_k$ converge absolutamente; es decir, que $\sum_{k=1}^{\infty} \|g_k\|$ converge. Entonces cualquier reordenación de la serie $\sum_{n=1}^{\infty} g_k$ también converge absolutamente, y al mismo límite.*

Solución Sea $\sum_{k=1}^{\infty} g_{\sigma(k)}$ la serie reordenada. Dado que $\varepsilon > 0$, existe N tal que $n \geq N$ implique

$$\|g_n\| + \cdots + \|g_{n+p}\| < \varepsilon.$$

Elíjase un entero N_1 tal que $\sigma(n) > N$ si $n > N_1$ (podemos hacer esto debido a que sólo existe un número finito de enteros n tales que $\sigma(n) \leq N$, pues σ es una biyección). Así,

si $n > N_1$, tenemos que $\sigma(n + k) > N$, de modo que

$$\|g_{\sigma(n)}\| + \cdots + \|g_{\sigma(n+p)}\| < \varepsilon.$$

Por el criterio de Cauchy, $\sum_{k=1}^{\infty} g_{\sigma(n)}$ converge absolutamente.

Para mostrar que los límites son iguales, dado ε elegimos $N_2 > N$, donde N es como antes, tal que si $1 \leq l \leq N$, entonces $l = \sigma(n)$ para algún n, $1 \leq n \leq N_2$. Esto se debe a que tales l son un número finito y σ es suprayectiva. Sea $N_0 = \text{máx}(N_1, N_2)$; entonces, para $m > N_0$,

$$\left\|\sum_{k=1}^{m} g_{\sigma(k)} - \sum_{n=1}^{\infty} g_n\right\| = \left\|\sum_{k=1}^{m} g_{\sigma(k)} - \sum_{n=1}^{N_0} g_n - \sum_{n=N_0+1}^{\infty} g_n\right\|$$

$$\leq \left\|\sum_{k=1}^{m} g_{\sigma(k)} - \sum_{n=1}^{N_0} g_n\right\| + \left\|\sum_{n=N_0+1}^{\infty} g_n\right\|$$

Por construcción, $\sum_{k=1}^{m} g_{\sigma(k)} - \sum_{n=1}^{N_0} g_n$ es una suma de *algunos* términos g_l, donde $l > N$. Entonces, los términos de la ecuación anterior son menores que $\varepsilon + \varepsilon = 2\varepsilon$.

Así, la serie $\sum_{k=1}^{m} g_{\sigma(k)}$ converge a $\sum_{n=0}^{\infty} g_n$, que es la conclusión deseada. El resultado de este ejemplo está íntimamente ligado con importantes teoremas de reordenaciones para series dobles (vease el ejercicio 51). ♦

Ejercicios del capítulo 5

1. **a.** Sea f_k una sucesión de funciones de $A \subset \mathbb{R}^n$ a $\mathbb{R}^m$. Supóngase que existen constantes m_k tales que $\|f_k(x) - f(x)\| \leq m_k$ para todo $x \in A$ y tales que $m_k \to 0$. Demuéstrese que $f_k \to f$ uniformemente.

b. Si $m_k \to m \in \mathbb{R}$ y $\|f_k(x) - f_l(x)\| \leq |m_k - m_l|$ para todo $x \in A$, muéstrese que f_k converge uniformemente.

2. Determínense cuáles de las siguientes *sucesiones* convergen (puntual o uniformemente) cuando $k \to \infty$. Verifíquese la continuidad del límite en cada caso.

a. $(\operatorname{sen} x)/k$ en $\mathbb{R}$

b. $1/(kx + 1)$ en $]0, 1[$

c. $x/(kx + 1)$ en $]0, 1[$

d. $x/(1 + kx^2)$ en $\mathbb{R}$

e. $(1, (\cos x)/k^2)$, una sucesión de funciones de $\mathbb{R}$ en $\mathbb{R}^2$

3. Determínense cuáles de las siguientes *series* de números reales $\sum_{k=1}^{\infty} g_k$ convergen (puntual o uniformemente). Verifíquese la continuidad del límite en cada caso.

a. $g_k(x) = \begin{cases} 0, & x \leq k \\ (-1)^k, & x > k. \end{cases}$

b. $g_k(x) = \begin{cases} 1/k^2, & |x| \leq k \\ 1/x^2, & |x| > k. \end{cases}$

c. $g_k(x) = \left(\frac{(-1)^k}{\sqrt{k}}\right) \cos(kx)$ en $\mathbb{R}$.

d. $g_k(x) = x^k$ en $]0, 1[$.

4. Sean $f_n : [1, 2] \to \mathbb{R}$ dadas por $f_n(x) = x/(1 + x)^n$.

a. Demuéstrese que $\sum_{n=1}^{\infty} f_n(x)$ es convergente para $x \in [1, 2]$.

b. ¿Es uniformemente convergente?

c. ¿Se cumple que $\int_1^2 \left(\sum_1^{\infty} f_n(x)\right)\, dx = \sum_1^{\infty} \int_1^2 f_n(x)\, dx$?

5. Supóngase que $f_k \to f$ uniformemente, donde $f_k : A \subset \mathbb{R}^n \to \mathbb{R}$, y que $g_k \to g$ uniformemente, donde $g_k : A \to \mathbb{R}$; existe M_1 constante tal que $\|g(x)\| \leq M_1$ para todo x y existe M_2 constante tal que $\|f(x)\| \leq M_2$ para todo x. Muéstrese que $f_k g_k \to fg$ uniformemente. Encuéntrese un contraejemplo si M_1 o M_2 no existen. ¿Se necesitan M_1 y M_2 para la convergencia puntual?

6. Demuéstrese que la sucesión $f_k : A \to \mathbb{R}^m$ converge puntualmente sii para cada $x \in A$, $f_k(x)$ es una sucesión de Cauchy.

7. Para las funciones $f : A \subset \mathbb{R}^n \to \mathbb{R}$, defínase C_b como en el texto. Muéstrese que $\|fg\| \leq \|f\|\,\|g\|$.

8. ¿Es cierto que la convergencia uniforme de funciones continuas en un conjunto compacto a un límite continuo implica la convergencia uniforme en ese conjunto?

9. Supóngase que las funciones g_k son continuas y que $\sum_{k=1}^{\infty} g_k$ converge uniformemente en $A \subset \mathbb{R}^n$. Si $x_k \to x_0$ en A, demuéstrese que $\sum_{n=1}^{\infty} g_n(x_k) \to \sum_{n=1}^{\infty} g_n(x_0)$ cuando $k \to \infty$.

10. Para las sucesiones y series de los ejercicios 2 y 3, ¿en qué casos podemos integrar o derivar término a término?

11. **a.** ¿Debe una contracción en cualquier espacio métrico tener un punto fijo? Razónese la respuesta.

b. Sea $f: X \to X$, donde X es un espacio métrico completo (como $\mathbb{R}$), una función que satisface $d(f(x), f(y)) < d(x, y)$ para todo $x, y \in X$. ¿Debe f tener un punto fijo? ¿Qué ocurre si X es compacto?

12. Una función $f: A \to \mathbb{R}$, donde $A \subset \mathbb{R}^n$, es ***semicontinua inferiormente*** si siempre que $x_0 \in A$ y $\lambda < f(x_0)$, existe una vecindad U de x_0 tal que $\lambda < f(x)$ para todo $x \in U \cap A$. La ***semicontinuidad superior*** se define de forma análoga.

a. Muéstrese que f es continua sii es semicontinua inferior y superiormente.

b. Si las funciones f_k son semicontinuas inferiormente, $f_k \to f$ puntualmente y $f_{k+1}(x) \geq f_k(x)$, demuéstrese entonces que f es semicontinua inferiormente.

c. En **b**, muéstrese que f no tiene por qué ser continua, aunque las funciones f_k lo sean.

d. Sea $f: [0,1] \to \mathbb{R}$ y $g(x) = \sup_{\delta>0} \inf_{|y-x|<\delta} f(y)$. Demuéstrese que g es semicontinua inferiormente.

13. En el teorema **5.3.3**. muéstrese que $f_k \to f$ uniformemente. [Sugerencia: utilícese el teorema del valor medio.]

14. Sea $f: X \to X$ una contracción en un espacio métrico compacto X. Muéstrese que $\bigcap_{i=1}^{\infty} f^n(X)$ es un único punto, donde $f_n = f \circ f \circ \cdots \circ f$ (n veces). ¿Es esto cierto si $X = \mathbb{R}$?

15. Sea $g_k \in \mathbb{R}^n$ y f_k una subsucesión de g_k. Demuéstrese que si $\sum g_k$ converge absolutamente, entonces $\sum f_k$ también converge absolutamente. Encuéntrese un contraejemplo si $\sum g_k$ sólo es convergente.

16. Obsérvese que en el ejemplo resuelto **5.5**, el mismo argumento se aplica en cualquier espacio normado. Utilícese esta observación y el espacio C_b para demostrar el siguiente:

Teorema *Sean $g_k : A \subset \mathbb{R}^n \to \mathbb{R}^m$ acotadas y continuas, y supóngase que $\sum g_k$ converge uniforme y absolutamente. Entonces cualquier reordenación converge también uniforme y absolutamente, y al mismo límite.*

17. Sea $\sum_{n=0}^{\infty} a_n$ una serie real convergente, pero no absolutamente. Dado cualquier número x, muéstrese que existe una reordenación $\sum b_n$ de esta serie que converge a x.

18. Encuéntrese un ejemplo de una sucesión de funciones discontinuas f_k que convergen *uniformemente* a una función límite f que es continua.

19. Demuéstrese que

$$\sum_{n=1}^{\infty} \left(\frac{\operatorname{sen} nx}{n^2}\right) x^3$$

define una función continua en todo $\mathbb{R}$.

20. Constrúyase la función $g(x)$ como $g(x) = |x|$ si $x \in [-1/2, 1/2]$ y extiéndase g de modo que sea periódica. Defínase

$$f(x) = \sum_{n=1}^{\infty} \frac{g(4^{n-1}x)}{4^{n-1}}.$$

a. Dibújense g y los primeros términos de la suma.

b. Mediante el criterio M de Weierstrass muéstrese que f es continua.

c. Demuéstrese que f no es derivable en *ningún punto*.

21. **a.** Demuéstrese que si $A \subset \mathbb{R}^n$ es compacto, $\mathcal{B} \subset C(A, \mathbb{R}^m)$ es compacto $\Leftrightarrow$ $\mathcal{B}$ es cerrado, acotado y equicontinuo.

b. Sea $D = \{ f \in C([0, 1], \mathbb{R}) \mid \|f\| \leq 1 \}$. Muéstrese que D es cerrado y acotado, pero *no* compacto. [Sugerencia: constrúyase una sucesión en D que no sea equicontinua y utilícese **a.**]

22. Sean $\mathcal{B} \subset C(A, \mathbb{R}^m)$ y $A \subset \mathbb{R}^m$ compacto. Supóngase que para cada $x_0 \in A$ y $\varepsilon > 0$, existe $\delta > 0$ tal que $d(x, x_0) < \delta$ implica $d(f(x), f(x_0)) < \varepsilon$ para todo $f \in \mathcal{B}$. Demuéstrese que $\mathcal{B}$ es equicontinuo.

23. Sea $f : \mathbb{R} \to \mathbb{R}$ y supóngase que $f \circ f$ es continua. ¿Debe f ser continua?

24. Un espacio métrico M verifica el ***segundo axioma de numerabilidad*** si existe una colección numerable $U_1, U_2, \ldots$ de conjuntos abiertos en M tales que cualquier conjunto abierto en M es la unión de elementos de esta colección. Demuéstrese que tal M tiene un subconjunto numerable C tal que $\mathrm{cl}(C) = M$ (decimos entonces que M es ***separable***). Recíprocamente, demuéstrese que un espacio métrico separable verifica el segundo axioma de numerabilidad.

25. Sea $g : [0, 1] \to \mathbb{R}$ continua e inyectiva. Muéstrese que g es o bien creciente o bien decreciente.

26. Sea $k(x, y)$ una función continua de valores reales sobre el cuadrado $U = \{(x, y) \mid 0 \leq x \leq 1, 0 \leq y \leq 1\}$ y supóngase que $|k(x, y)| < 1$ para cada $(x, y) \in U$. Sea $A : [0, 1] \to \mathbb{R}$ continua. Demuéstrese que existe una única función continua de valores reales $f(x)$ en $[0, 1]$ tal que

$$f(x) = A(x) + \int_0^1 k(x, y) f(y)\, dy.$$

27. Sea $f :]a, b[\to \mathbb{R}$ uniformemente continua y supóngase que $x_n \to b$. Muéstrese que $\lim_{n\to\infty} f(x_n)$ existe.

28. Sea $f_n(x) = x/n$. ¿Es f_n uniformemente convergente en $[0, 396]$? ¿Y en $\mathbb{R}$?

29. Analícese la continuidad uniforme de las siguientes funciones:

a. $f(x) = x^2, \quad x \in \,]-1, 1[.$

b. $f(x) = x^{1/3}, \quad x \in [0, \infty[.$

c. $f(x) = e^{-x}, \quad x \in [0, \infty[.$

d. $f(x) = \dfrac{x \operatorname{sen} 1}{x}, \quad 0 < x \leq 1, f(0) = 0.$

e. $f(x) = \operatorname{sen}[\ln(1 + x^3)], \quad -1 < x \leq 1, f(-1) = 0.$

30. Sea $B = \{ f \in C([0, 1], \mathbb{R}) \mid f$ es C^1 en $]0, 1[, f(0) = 0$ y $|f'(x)| \leq 1\}$. Muéstrese que $\mathrm{cl}(B)$ es compacto.

31. Sea a_n una sucesión convergente de números reales, $a_n \to a$. Sea $b_n = (a_1 + \cdots + a_n)/n$. Muéstrese que también $b_n \to a$.

32. Sean M y N espacios métricos y $f: M \to N$ continua. Supóngase que $f(M)$ está formado por dos puntos distintos. Demuéstrese que M no es conexo.

33. Sea $f_n : [0, 1] \to \mathbb{R}$ una sucesión de funciones crecientes en $[0, 1]$ y supóngase que $f_n \to 0$ puntualmente. ¿Debe f_n converger uniformemente? ¿Qué ocurre si f_n sólo converge puntualmente a algún límite f?

34. Encuéntrese una sucesión $f_n : [0, 1] \to \mathbb{R}$ de funciones derivables tales que $f_n \to 0$ uniformemente, pero tales que $f_n'(1/2)$ no converja a cero.

35. Sea $f: \mathbb{R} \to \mathbb{R}$ continua y biyectiva. Muéstrese que f^{-1} es continua. (Véase el ejercicio 7, capítulo 4. Para una generalización, léase M. Hoffman, "Continuity of Inverse Functions", *Mathematics Magazine* **48** (1975), págs. 66–73.)

36. Sea $f(x, y) = x^2y/(x^4 + y^2)$. Analícese el comportamiento de f cerca de $(0, 0)$ con respecto de los límites

a. $\lim_{(x,y)\to(0,0)} f(x, y),$

b. $\lim_{x\to 0}[\lim_{y\to 0} f(x, y)],$

c. $\lim_{y\to 0}[\lim_{x\to 0} f(x, y)].$

37. Supóngase que $f: \mathbb{R} \to \mathbb{R}$ es continua y que $f(1) = 7$. Supóngase además que $f(x)$ es racional para todo x. Demuéstrese que f es constante.

38. Demuéstrese que $1 + 1/2 + 1/4 + 1/8 + \cdots$ converge, y que $1 - 1/2 + 1/3 - 1/4 + \cdots$ converge, pero no absolutamente.

39. Una función $g : [0, 1] \to \mathbb{R}$ es ***simple*** si podemos dividir $[0, 1]$ en subintervalos donde g sea constante, excepto tal vez en los extremos. Sea $f: [0, 1] \to \mathbb{R}$ continua y $\varepsilon > 0$. Demuéstrese que existe una función simple g tal que $\|f - g\| < \varepsilon$.

40. **a.** Defínase $\delta : C([0, 1], \mathbb{R}) \to \mathbb{R}, f \mapsto f(0)$. Demuéstrese que δ es lineal y continua.

b. Sea $g : \mathbb{R} \to \mathbb{R}$ continua. Defínase $F : C([0, 1], \mathbb{R}) \to \mathbb{R}$, $C([0, 1], \mathbb{R})$ como $F(f) = g \circ f$. Demuéstrese que F es continua y demuéstrese también que si g es uniformemente continua, entonces F es uniformemente continua.

41. Muéstrese que existe un polinomio $p(x)$ tal que $|p(x) - |x|^3| < 1/10$ para $-1000 \leq x \leq 1000$.

42. Estúdiese la posibilidad de reemplazar la sucesión de polinomios de Bernstein en el teorema **5.8.1** por una sucesión de polinomios de interpolación de Lagrange (véase el ejercicio 2, §**5.8**, para su definición y propiedades) para realizar la demostración de los teoremas en §**5.8**.

43. Sea $C_e([-1, 1], \mathbb{R})$ el conjunto de las funciones pares en $C([-1, 1], \mathbb{R})$.

a. Muéstrese que C_e es cerrado y no es denso en C.

b. Muéstrese que los polinomios pares son densos en C_e, pero no en C.

44. *Proyectos:* Examine la posibilidad de extender el teorema de Stone-Weierstrass a:

a. Funciones de valores complejos (consérvense las mismas hipótesis sobre B, pero agréguese "$f \in B$ implica $\bar{f} \in B$", donde la barra indica la conjugación compleja).

b. Dominios no compactos (consúltese Simmons, *Introduction to Topology and Modern Analysis,* McGraw-Hill).

c. Utilícese **b** para estudiar la densidad de las funciones de Hermite en un espacio adecuado de funciones continuas (las funciones de Hermite se definen y analizan, por ejemplo, en Courant y Hilbert, *Methods of Mathematical Physics, vol. I*, Wiley).

45. **a.** Sea $f_n : K \subset \mathbb{R}^n \to \mathbb{R}^m$ una sucesión de funciones equicontinuas en un conjunto compacto K que convergen puntualmente. Demuéstrese que la convergencia es uniforme.

b. Sea

$$f_n(x) = \frac{x^2}{x^2 + (1 - nx)^2}, \quad 0 \leq x \leq 1.$$

Muéstrese que f_n converge puntualmente, pero no uniformemente. ¿Qué se puede concluir de **a**?

46. Sea $f(t, x)$ definida y continua en $a \leq t \leq b$ y $x \in \mathbb{R}^n$. El propósito de este ejercicio es mostrar que el problema $dx/dt = f(t, x)$, $x(a) = x_0$ tiene una solución en un intervalo $t \in [a, c]$ para algún $c > a$ (es única sólo bajo condiciones más fuertes). Realícense las siguientes operaciones: divídase $[a, b]$ en n partes iguales $t_0 = a, \ldots, t_n = b$ y defínase una función continua x_n inductivamente mediante

$$\begin{cases} x_n'(t) = f(t_i, x_n(t_i)), & t_i < t < t_{i+1}, \\ x_n(a) = x_0. \end{cases}$$

Tómese $\Delta_n(t) = x_n'(t) - f(t, x_n(t))$, de modo que

$$x_n(t) = x_0 + \int_a^t f(s, x_n(s)) + \Delta_n(s)\, ds.$$

Usando el teorema de Arzela-Ascoli encuéntrese una subsucesión convergente de x_n. Muéstrese que el límite satisface $dx/dt = f(t, x)$ y $x(a) = x_0$. Este método se denomina ***aproximación poligonal.***

47. Sea $f_n : A \subset \mathbb{R}^n \to \mathbb{R}^m$ una sucesión de funciones equicontinuas. Supóngase que f_n converge en un subconjunto denso K de A. Demuéstrese que la sucesión converge en todo A. ¿Arroja esto alguna luz sobre la demostración del teorema **5.6.2**?

48. Demuéstrese que la norma del supremo en $C([0, 1], \mathbb{R})$ no se obtiene de un producto interno $\langle\ ,\ \rangle$ como $\|f\|^2 = \langle f, f\rangle$. [Sugerencia: muéstrese que falla la identidad del paralelogramo.]

49. Sea S un conjunto y sea $\mathcal{B}$ el conjunto de *todas* las funciones acotadas de valores reales en S; dótese a $\mathcal{B}$ de la norma del supremo. Demuéstrese que $\mathcal{B}$ es un espacio de Banach.

50. Sea $f: \mathbb{R} \to \mathbb{R}$ un límite uniforme de polinomios. Demuéstrese que f es un polinomio.

51. Considérese una serie doble

$$\sum_{m,n=0}^{\infty} a_{mn} \quad \text{donde } a_{mn} \in \mathbb{R}, \quad m, n = 0, 1, 2, \ldots.$$

Decimos que esta serie ***converge a*** S si para cada $\varepsilon > 0$ existe N tal que $n, m \geq N$ implica

$$\left| \sum_{k,l=0}^{m,n} a_{kl} - S \right| < \varepsilon.$$

Defínase la ***convergencia absoluta*** y demuéstrese que si $\sum_{m,n=0}^{\infty} a_{nm}$ es absolutamente convergente, entonces la suma se puede ***reordenar*** como sigue:

$$\sum_{m,n=0}^{\infty} a_{nm} = \sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} a_{nm}\right) = \sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} a_{nm}\right)$$

Interprétese este resultado en términos de la suma de entradas de una matriz con un número infinito de filas y columnas.

52. ¿Podemos derivar la serie

$$x = \sum_{k=1}^{\infty}\left(\frac{x^k}{k} - \frac{x^{k+1}}{k+1}\right), \quad 0 \le x \le 1$$

término a término?

53. Evalúense los siguientes límites:

a. $\lim_{x \to 0} \dfrac{1 - \cos x}{3^x - 2^x}$

b. $\lim_{x \to 0^+} (1 + \operatorname{sen} 2x)^{1/x}$

c. $\lim_{x \to 0^+} \left(\dfrac{1}{\operatorname{sen} x} - \dfrac{1}{x}\right)$

54. Verifíquese la convergencia o divergencia de las siguientes series infinitas:

a. $\displaystyle\sum_{k=1}^{\infty} \frac{\sqrt{k}\log k}{k^2 + 2k + 3}$

b. $\displaystyle\sum_{k=1}^{\infty} \frac{k!3^k}{k^k}$

c. $\sum_{k=1}^{\infty} \dfrac{(k!)^2}{(2k)!}$

55. Demuéstrese que $\pi/4 = 1 - 1/3 + 1/5 - 1/7 + \cdots$, partiendo de

$$(1 + x^2)^{-1} = \sum_{k=0}^{\infty}(-1)^k x^{2k}, \quad |x| < 1.$$

56. Verifíquese la convergencia absoluta y condicional de las siguientes series:

a. $\displaystyle\sum_{n=1}^{\infty}(-1)^n n^{-\alpha}$, α real

b. $\sum_{k=3}^{\infty} \frac{(-1)^k \log k}{k \log \log k}$

c. $\sum_{k=2}^{\infty} \frac{(-1)^k}{k^\alpha + (-1)^k}$, $\alpha > 0$

57. Demuéstrese que si

a. $f_n(x)$, $g(x)$ son continuas, $0 \leq x < \infty$,

b. $|f_n(x)| \leq g(x)$, $n = 1, 2, 3, \ldots$, $0 \leq x < \infty$,

c. $f_n(x) \to f(x)$ uniformemente, $0 \leq x \leq R$, para cualquier $R < \infty$, y

d. $\int_0^\infty g(x)\, dx < \infty$,

entonces $\lim_{n\to\infty} \int_0^\infty f_n(x)\, dx = \int_0^\infty f(x)\, dx$.

58. Demuéstrense los siguientes criterios de convergencia:

a. Si $u_n > 0$, $\frac{u_{n+1}}{u_n} \leq 1 - \frac{1}{n} - \frac{\alpha}{n \log n}$, $\alpha > 1$, entonces $\sum u_n$ converge.

b. Si $u_n > 0$, $\frac{u_{n+1}}{u_n} \geq 1 - \frac{1}{n} - \frac{1}{n \log n}$, entonces $\sum u_n$ diverge.

59. **a.** Sea $p > 1$ con $1/p + 1/q = 1$. Para a, b, $t > 0$, demuéstrese que

$$ab \leq \frac{a^p t^p}{p} + \frac{b^q t^{-q}}{q}$$

y que ab es el valor mínimo del miembro derecho (una forma de demostrar esto es usando el cálculo elemental).

b. Demuéstrese la ***desigualdad de Hölder***: si $a_k, b_k \geq 0$, $p > 1$ y $1/p + 1/q = 1$, entonces

$$\sum_1^n a_k b_k \leq \left(\sum_1^n a_k^p\right)^{1/p} \left(\sum_1^n b_k^p\right)^{1/q}$$

[Sugerencia: imítese la demostración de la desigualdad de Cauchy-Schwarz, usando el apartado **a.**]

c. Demuéstrese la ***desigualdad de Minkowski***: si $a_k, b_k \geq 0$ y $p > 1$, entonces

$$\left(\sum_1^n (a_k + b_k)^p\right)^{1/p} \leq \left(\sum_1^n a_k^p\right)^{1/p} + \left(\sum_1^n b_k^p\right)^{1/p}$$

60. La serie $\sum_{k=1}^\infty x^k/k$ converge si $|x| < 1$. ¿Para qué números complejos x tales que $|x| = 1$ converge la serie?

61. Sea $\sum a_k x^k$ una serie con radio de convergencia R. Muéstrese que $\sum a_k(x-b)^k$ converge dentro del disco de centro b y radio R.

62. Determínese el radio de convergencia de

a. $\sum x^k k/(k+1)$

b. $\sum x^k/\log k$

63. (Serie binómica) Considérese

$$\sum_{k=0}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}x^k.$$

Supóngase que α no es un entero mayor o igual que 0. Muéstrese que el radio de convergencia es $R = 1$ (véase el ejercicio 49, capítulo 2, acerca de la serie hipergeométrica, para el comportamiento de la serie en $x = \pm 1$).

64. ¿La serie $1 + 1/2 + 1/3 + \cdots$ converge $(C, 1)$ o (Abel)?

65. Sea $f(x)$ continua, $0 \leq x < \infty$. Por lo general, definimos

$$\int_0^{\infty} f(x)\,dx = \lim_{R\to+\infty} \int_0^R f(x)\,dx,$$

si el límite existe. Por analogía con la sumabilidad $(C, 1)$, defínase un concepto de "***integrabilidad*** $(C, 1)$ de 0 a ∞" y demuéstrese que tal método de integrabilidad es regular; esto es, que coincide con el usual $\int_0^{\infty}$ si este último converge.

66. Defínase t_n inductivamente como $t_1 = 1$ y $t_{n+1} = t_n/(1 + t_n^{\beta})$, donde β está fijo, $0 \leq \beta < 1$. Demuéstrese que $\sum_{k=1}^{\infty} t_n$ converge. [Sugerencia: encuéntrese C constante tal que $t_n \leq C/n^{1/\beta}$.]

67. Sean $A = \{ j/2^n \in [0, 1] \mid n = 1, 2, 3, \ldots, j = 0, 1, 2, \ldots, 2^n\}$ y $f: A \to \mathbb{R}$ una función que satisface la siguiente condición: existe una sucesión $\varepsilon_n > 0$ con $\sum_{n=1}^{\infty} \varepsilon_n < \infty$ y

$$\left| f\left(\frac{j-1}{2^n}\right) - f\left(\frac{j}{2^n}\right) \right| < \varepsilon_n \quad \text{para todo } n > 0,\ j = 1, 2, \ldots, 2^n.$$

Demuéstrese que f tiene una única extensión a una función continua de $[0, 1]$ a $\mathbb{R}$.

68. Sean $A \subset \mathbb{R}^n$ compacto y $\mathcal{B} \subset C(A, \mathbb{R}^n)$ compacto. Demuéstrese que $\mathcal{B}$ es equicontinuo, como sigue:

a. Demuéstrese que la transformación $E : C(A, \mathbb{R}^m) \times A \to \mathbb{R}^m$ definida por $(f, x) \mapsto f(x)$ es continua.

b. Utilícese la continuidad uniforme de E restringida a $\mathcal{B} \times A$ para deducir el resultado (este método de demostración se debe a J. Allen).

69. Sean las funciones f_n monótonas crecientes y continuas en $[0, 1]$. Supóngase que $F(x) = \sum_{n=1}^{\infty} f_n(x)$ converge para cada $x \in [0, 1]$. Demuéstrese que F es continua.

Capítulo 6
Transformaciones diferenciables

En este capítulo desarrollamos el concepto de transformación diferenciable de $\mathbb{R}^n$ en $\mathbb{R}^m$ (o, más en general, de un espacio normado a otro). A partir de este capítulo usaremos algo de álgebra lineal. En este momento, convendría que el lector repase el concepto de transformación lineal y su representación matricial, pues definiremos la derivada como una transformación lineal. La conexión entre la definición general y las derivadas parciales aparece en §**6.2**. Después de esto generalizamos los teoremas usuales del cálculo del contexto de una variable, estudiados en §**4.7**, al caso de varias variables (como el teorema que establece que la diferenciabilidad implica la continuidad, la regla de la cadena, el teorema del valor medio, el teorema de Taylor, los criterios para los extremos, etcétera).

§6.1 Definición de derivada

Una función de una variable $f\colon]a, b[\to \mathbb{R}$ es ***derivable (o diferenciable) en*** $x_0 \in]a, b[$ si el límite

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

existe. De forma equivalente, podemos escribir esta fórmula como

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - f'(x_0)h}{h} = 0,$$

es decir,

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} = 0$$

o, lo que es lo mismo,

$$\lim_{x \to x_0} \frac{|f(x) - f(x_0) - f'(x_0)(x - x_0)|}{|x - x_0|} = 0.$$

El número $f'(x_0)$ representa la pendiente de la recta tangente a la gráfica de f en el punto $(x_0, f(x_0))$, como en la figura 6.1-1.

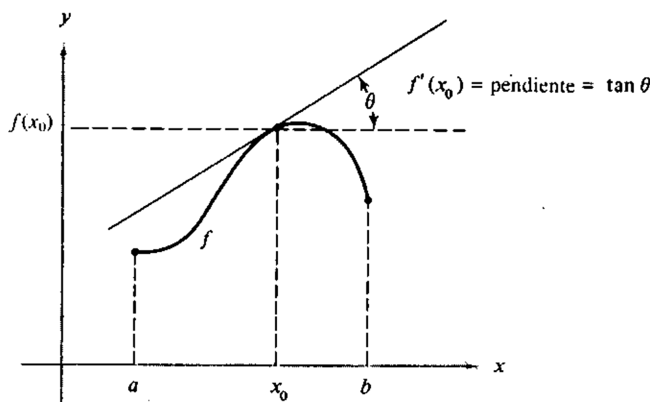


FIGURA 6.1-1 La derivada es la pendiente de la recta tangente

La derivabilidad (o diferenciabilidad) de $f : \mathbb{R} \to \mathbb{R}$ en x_0 es equivalente a la existencia de un número m tal que

$$\lim_{x \to x_0} \frac{|f(x) - f(x_0) - m(x - x_0)|}{|x - x_0|} = 0.$$

La función $T(s) = ms$ es lineal: $T(\alpha s + \beta t) = \alpha T(s) + \beta T(t)$ para todo α, β, s y t reales. Para generalizar este concepto al caso de las transformaciones $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, damos la siguiente definición.

6.1.1 Definición *Una transformación $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ es* ***diferenciable en*** *$x_0 \in A$ si existe una transformación lineal, denotada $\mathbf{D}f(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ y llamada* ***diferencial*** *de f en x_0 tal que*

$$\lim_{x \to x_0} \frac{\|f(x) - f(x_0) - \mathbf{D}f(x_0)(x - x_0)\|}{\|x - x_0\|} = 0.$$

En esta definición, $\mathbf{D}f(x_0)(x - x_0)$ denota el valor de la transformación lineal $\mathbf{D}f(x_0)$ aplicada al vector $x - x_0 \in \mathbb{R}^n$, de modo que $\mathbf{D}f(x_0)(x - x_0) \in \mathbb{R}^m$. Con frecuencia escribiremos $\mathbf{D}f(x_0) \cdot h$ en vez de $\mathbf{D}f(x_0)(h)$.

Nos concentraremos en el caso euclídeo, pero debemos observar ahora que si V, W son espacios normados y $f : A \subset V \to W$, entonces $\mathbf{D}f(x_0) : V \to W$ y la definición anterior tiene sentido.

Podemos reescribir la definición diciendo que para cada $\varepsilon > 0$ existe $\delta > 0$ tal que $x \in A$ y $\|x - x_0\| < \delta$ implican

$$\|f(x) - f(x_0) - \mathbf{D}f(x_0)(x - x_0)\| \leq \varepsilon\|x - x_0\|.$$

En **6.1.1**, está implícito que $x \neq x_0$, pero en esta formulación $\varepsilon - \delta$, podemos permitir que $x = x_0$, pues ambos miembros se reducen a cero.

Intuitivamente, se supone que $x \mapsto f(x_0) + \mathbf{D}f(x_0)(x - x_0)$ es la *mejor aproximación afín* a f cerca del punto x_0 (una transformación afín es una transformación lineal más una constante). Véase la figura 6.1-2. En esta figura hemos indicado las ecuaciones de los planos tangentes a la gráfica de f.

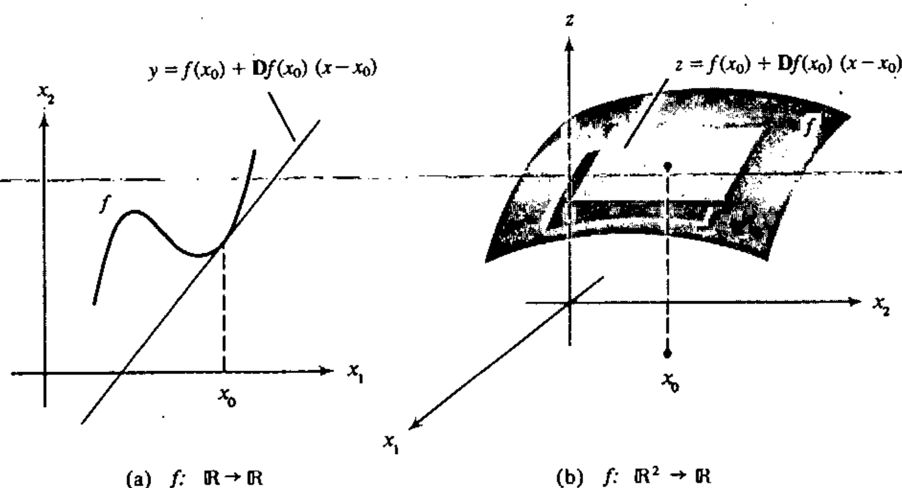


FIGURA 6.1-2 (a) $f : \mathbb{R} \to \mathbb{R}$; (b) $f : \mathbb{R}^2 \to \mathbb{R}$

Si f es una función de una variable y comparamos las definiciones de $\mathbf{D}f(x)$ y $df/dx = f'(x)$, vemos que $\mathbf{D}f(x)(h) = f'(x) \cdot h$ (el producto de los números $f'(x)$ y $h \in \mathbb{R}$). Así, la transformación lineal $\mathbf{D}f(x)$ es simplemente la multiplicación por df/dx, en este caso. Si f es diferenciable en cada punto de A, decimos que f es ***diferenciable en*** A. Esperamos, intuitivamente (como en la figura 6.1-2) que sólo puede haber una mejor aproximación lineal, lo cual es cierto si suponemos que A es un conjunto abierto.

6.1.2 Teorema *Sea A un conjunto abierto en $\mathbb{R}^n$ y supóngase que $f: A \to \mathbb{R}^m$ es diferenciable en x_0. Entonces $\mathbf{D}f(x_0)$ queda determinada de forma única por f.*

6.1.3 Ejemplo *Sea $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$. Calcúlense $\mathbf{D}f(x)$ y df/dx.*

Solución Por el cálculo de una variable, $dx^3/dx = 3x^2$. Así, en este ejemplo, $\mathbf{D}f(x)$ es la transformación lineal

$$h \mapsto \mathbf{D}f(x) \cdot h = 3x^2 h. \quad \blacklozenge$$

6.1.4 Ejemplo *Muéstrese que, en general, $\mathbf{D}f$ no está determinada de forma única.*

Solución Por ejemplo, si $A = \{x_0\}$ es un único punto, cualquier $\mathbf{D}f(x_0)$ satisface la condición, pues $x \in A$ tal que $\|x - x_0\| < \delta$ se cumple sólo si $x = x_0$, en cuyo caso la expresión

$$\|f(x) - f(x_0) - \mathbf{D}f(x_0)(x - x_0)\|$$

se anula. La definición se cumple entonces de manera trivial. $\blacklozenge$

Nota. Si se examina con detalle la demostración del teorema **6.1.2**, veremos que $\mathbf{D}f(x)$ es única (si existe) en un rango más amplio de conjuntos que los conjuntos abiertos. Por ejemplo, el teorema es válido para intervalos cerrados en $\mathbb{R}$ o, en general, para discos cerrados en $\mathbb{R}^n$.

Ejercicios de §6.1

1. Calcúlese $\mathbf{D}f(x)$ para $f: \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x \operatorname{sen} x$.

2. Demuéstrese que $\mathbf{D}(f + g) = \mathbf{D}f + \mathbf{D}g$.

3. Sea $A = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1,\ y = 0\}$. Demuéstrese que la conclusión del teorema **6.1.2** es falsa para este conjunto A.

4. Sea $f: \mathbb{R}^n \to \mathbb{R}^m$ y supóngase que existe M constante tal que para $x \in \mathbb{R}^n$, $\|f(x)\| \leq M\|x\|^2$. Demuéstrese que f es diferenciable en $x_0 = 0$ y que $\mathbf{D}f(x_0) = 0$.

5. Si $f: \mathbb{R} \to \mathbb{R}$ es diferenciable y $|f(x)| \leq |x|$, ¿debe ocurrir que $\mathbf{D}f(0) = 0$?

§6.2 Representación matricial

Además de la definición de $\mathbf{D}f(x)$, hay otra forma de derivar una función f de varias variables. Escribimos f en componentes, $f(x_1, \ldots, x_n) = f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)$, y calculamos las derivadas parciales $\partial f_j / \partial x_i$ para $j = 1, \ldots, m$ e $i = 1, \ldots, n$, donde la notación $\partial f_j / \partial x_i$ indica que calculamos la derivada usual de f_j con respecto de x_i mientras mantenemos las demás variables $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ fijas.

6.2.1 Definición *$\partial f_j / \partial x_i$ está dada por el siguiente límite, cuando éste existe:*

$$\frac{\partial f_j}{\partial x_i}(x_1, \ldots, x_n) = \lim_{h \to 0} \left\{ \frac{f_j(x_1, \ldots, x_i + h, \ldots, x_n) - f_j(x_1, \ldots, x_n)}{h} \right\}.$$

En §**6.1** vimos que $\mathbf{D}f(x)$ para $f : \mathbb{R} \to \mathbb{R}$ es simplemente la transformación lineal "multiplicación por df/dx". Podemos generalizar este hecho en el siguiente teorema.

6.2.2 Teorema *Supóngase que $A \subset \mathbb{R}^n$ es un conjunto abierto y que $f : A \to \mathbb{R}^m$ es diferenciable en A. Entonces las derivadas parciales $\partial f_j / \partial x_i$ existen y la matriz de la transformación lineal $\mathbf{D}f(x)$ con respecto de las bases canónicas en $\mathbb{R}^n$ y $\mathbb{R}^m$ está dada por*

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots & \dfrac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \dfrac{\partial f_m}{\partial x_2} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix},$$

*donde cada derivada parcial se evalúa en $x = (x_1, \ldots, x_n)$. Esta matriz es la **matriz jacobiana de f o matriz derivada**.*

Al hacer cálculos específicos, en general podemos calcular fácilmente la matriz jacobiana y el teorema **6.2.2** nos da $\mathbf{D}f$. En algunos textos, se llama a $\mathbf{D}f$ *la derivada* o *derivada total* de f.

Cuando $m = 1$, f es una función de n variables escalar. Entonces $\mathbf{D}f$ tiene la matriz

$$\left(\frac{\partial f}{\partial x_1} \cdots \frac{\partial f}{\partial x_n}\right)$$

y la diferencial aplicada a un vector $e = (a_1, \ldots, a_n)$ es

$$\mathbf{D}f(x) \cdot e = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} a_i.$$

Debemos enfatizar que $\mathbf{D}f$ es una transformación lineal en cada $x \in A$ y *la definición de* $\mathbf{D}f(x)$ *es independiente de la base utilizada.* Si cambiamos de la base canónica a otra base, los elementos de la matriz también se modifican. Si examinamos la definición de la matriz de una transformación lineal, se puede ver que las columnas de la matriz con respecto a la nueva base serán la diferencial $\mathbf{D}f(x)$ aplicada a la nueva base de $\mathbb{R}^n$, donde el vector imagen se expresa en la nueva base de $\mathbb{R}^m$. Por supuesto, *la propia transformación lineal* $\mathbf{D}f(x)$ *no cambia de base a base.* En el caso $m = 1$, $\mathbf{D}f(x)$ es, en la base canónica, una matriz $1 \times n$, como acabamos de mostrar. El vector cuyas componentes son iguales a las de $\mathbf{D}f(x)$ se denomina ***gradiente*** de f y se denota grad f o ∇f. Así, para cada $f: A \subset \mathbb{R}^n \to \mathbb{R}$,

$$\operatorname{grad} f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right).$$

(¡A veces se dice que grad f es simplemente $\mathbf{D}f$ con comas insertadas!) Para dar una definición intrínseca del gradiente debemos usar el producto interno. De hecho, $\nabla f(x)$ es el vector tal que para cualquier $v \in \mathbb{R}^n$,

$$\langle \nabla f(x), v \rangle = \mathbf{D}f(x) \cdot v.$$

Un caso particular importante ocurre cuando $f = L$ ya es lineal. Por la definición de la diferencial, vemos que $\mathbf{D}L = L$, como se esperaba, ya que la mejor aproximación afín a una transformación lineal tiene a la propia transformación lineal como parte lineal (véase el ejemplo **6.2.4**). Así, la matriz jacobiana de L es la propia matriz de L en este caso. Otro caso interesante es el de una transformación constante. De hecho, se ve que una transformación constante tiene diferencial cero, donde cero es la transformación lineal $K: \mathbb{R}^n \to \mathbb{R}^m$ tal que $K(x) = 0 = (0, \ldots, 0)$ para todo $x \in \mathbb{R}^n$.

6.2.3 Ejemplo *Sea* $f: \mathbb{R}^2 \to \mathbb{R}^3$ *dada por* $f(x, y) = (x^2, x^3 y, x^4 y^2)$. *Calcúlese* $\mathbf{D}f$.

Solución Según el teorema **6.2.2**, $\mathbf{D}f(x, y)$ es la transformación lineal cuya matriz es

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \\ \dfrac{\partial f_3}{\partial x} & \dfrac{\partial f_3}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x & 0 \\ 3x^2y & x^3 \\ 4x^3y^2 & 2x^4y \end{pmatrix}$$

donde $f_1(x, y) = x^2$, $f_2(x, y) = x^3y$, y $f_3(x, y) = x^4y^2$. ♦

6.2.4 Ejemplo *Sea* $L : \mathbb{R}^n \to \mathbb{R}^m$ *una transformación lineal; es decir,* $L(x + y) = L(x) + L(y)$ *y* $L(\alpha x) = \alpha L(x)$. *Muéstrese que* $\mathbf{D}L(x) = L$.

Solución Dados x_0 y $\varepsilon > 0$, debemos encontrar $\delta > 0$ tal que $\|x - x_0\| < \delta$ implique

$$\|L(x) - L(x_0) - \mathbf{D}L(x) \cdot (x - x_0)\| \leq \varepsilon\|x - x_0\|.$$

Pero si $\mathbf{D}L(x) = L$, el miembro izquierdo resulta

$$\|L(x) - L(x_0) - L(x - x_0)\|,$$

que se anula, pues $L(x - x_0) = L(x) - L(x_0)$, por la linealidad de L. Por lo tanto, $\mathbf{D}L(x) = L$ satisface la definición de la derivada (para cualquier $\delta > 0$). ♦

6.2.5 Ejemplo *Sea* $f(x, y, z) = (x \operatorname{sen} y)/z$. *Calcúlese* grad f.

Solución grad $f = (\partial f/\partial x, \partial f/\partial y, \partial f/\partial z)$, y

$$\frac{\partial f}{\partial x} = \frac{\operatorname{sen} y}{z}, \quad \frac{\partial f}{\partial y} = \frac{x \cos y}{z}, \quad \frac{\partial f}{\partial z} = -\frac{x \operatorname{sen} y}{z^2},$$

de modo que

$$\operatorname{grad} f(x, y, z) = \left(\frac{\operatorname{sen} y}{z}, \frac{x \cos y}{z}, -\frac{x \operatorname{sen} y}{z^2}\right)$$ ♦

Ejercicios de §6.2

1. Sea $f : \mathbb{R}^3 \to \mathbb{R}^2, f(x, y, z) = (x^4 y, xe^z)$. Calcúlese $\mathbf{D}f$.
2. Sea $f : \mathbb{R}^3 \to \mathbb{R}, (x, y, z) \mapsto e^{x^2+y^2+z^2}$. Calcúlense $\mathbf{D}f$ y grad f.
3. Sean L una transformación lineal de $\mathbb{R}^n \to \mathbb{R}^m$, $g : \mathbb{R}^n \to \mathbb{R}^m$ tal que $\|g(x)\| \leq M \|x\|^2$, y sea $f(x) = L(x) + g(x)$. Demuéstrese que $\mathbf{D}f(0) = L$.
4. Sea $f(x, y) = (xy, y/x)$. Calcúlese $\mathbf{D}f$. Calcúlese la matriz de $\mathbf{D}f(x, y)$ con respecto de la base $(1, 0), (1, 1)$ de $\mathbb{R}^2$.
5. Analícese la posibilidad de definir $\mathbf{D}f$ para f una transformación de un espacio normado en otro.

§6.3 Continuidad de las transformaciones diferenciables; curvas diferenciables

En §**4.7** se mostró que una función derivable de una variable es continua. Esto es claro desde un punto de vista intuitivo, pues el hecho de tener una recta tangente (o un plano tangente) a la gráfica es más fuerte que no tener saltos en la misma. Recordemos la demostración: Sea $f :]a, b[\to \mathbb{R}$ derivable en x_0. Entonces

$$\begin{aligned}\lim_{x \to x_0} (f(x) - f(x_0)) &= \lim_{x \to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0}\right) \cdot (x - x_0)\\ &= f'(x_0) \cdot \lim_{x \to x_0} (x - x_0) = f'(x_0) \cdot 0 = 0\end{aligned}$$

de modo que $\lim_{x \to x_0} (f(x) - f(x_0)) = 0$, lo que implica que f es continua en x_0. El argumento se generaliza al caso de $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, como en la siguiente proposición.

6.3.1 Proposición *Supóngase que $A \subset \mathbb{R}^n$ es abierto y que $f: A \to \mathbb{R}^m$ es diferenciable en A. Entonces f es continua. De hecho, para cada $x_0 \in A$ existe una constante $M > 0$ y un $\delta_0 > 0$ tales que $\|x - x_0\| < \delta_0$ implique $\|f(x) - f(x_0)\| \leq M\|x - x_0\|$ (esto recibe el nombre de **propiedad local de Lipschitz**).*

Anteriormente hemos analizado el caso de las funciones de valores reales, $f : \mathbb{R}^n \to \mathbb{R}$. El caso de una función $c : \mathbb{R} \to \mathbb{R}^m$ también es importante; c representa aquí una ***curva*** o

arco en $\mathbb{R}^m$. La diferencial $\mathbf{D}c(t) : \mathbb{R} \to \mathbb{R}^m$ está representada por el vector asociado a la matriz con una columna

$$\begin{pmatrix} \dfrac{dc_1}{dt} \\ \vdots \\ \dfrac{dc_m}{dt} \end{pmatrix}$$

donde $c(t) = (c_1(t), \ldots, c_m(t))$. Este vector se denota $c'(t)$ y se llama el ***vector tangente*** o ***vector velocidad*** de la curva. Si observamos que

$$c'(t) = \lim_{h \to 0} \frac{c(t+h) - c(t)}{h}$$

y usamos el hecho de que $(c(t+h) - c(t))/h$ es una cuerda que aproxima la recta tangente a la curva, vemos que $c'(t)$ debe representar exactamente el vector tangente (véase la figura 6.3-1). En términos de una partícula en movimiento, $(c(t+h) - c(t))/h$ es una aproximación de la velocidad, pues es distancia/tiempo, de modo que $c'(t)$ es la ***velocidad instantánea***.

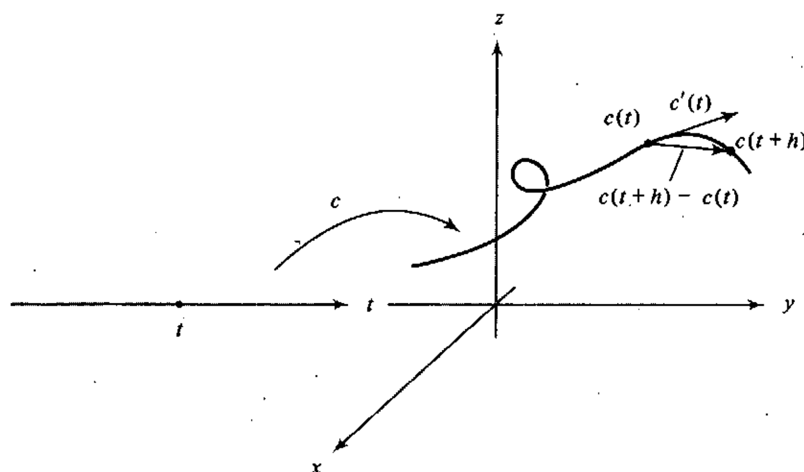


FIGURA 6.3-1 El vector velocidad de una curva

Estrictamente, siempre deberíamos representar $c'(t)$ como un vector columna, pues la matriz de $\mathbf{D}c(t)$ es una matriz $m \times 1$. Sin embargo, tipográficamente resulta muy complicado, de modo que solemos escribir $c'(t)$ como un vector fila.

6.3.2 Ejemplo *Sea $c(t) = (t^2, t, \text{sen } t)$. Determínese el vector tangente a $c(t)$ en $c(0) = (0, 0, 0)$.*

Solución $c'(t) = (2t, 1, \cos t)$. Si $t = 0$, $c'(0) = (0, 1, 1)$, que es el vector tangente a $c(t)$ en $(0, 0, 0)$. ♦

A continuación recordaremos algunas ideas de **§4.7** para el caso de una variable.

6.3.3 Ejemplo *Demuéstrese que $f : \mathbb{R} \to \mathbb{R}$ dada por $x \mapsto |x|$ es continua, pero no derivable en* 0.

Solución $f(x) = x$ para $x \geq 0$ y $f(x) = -x$ para $x < 0$, por lo que f es continua en $]0, \infty[$ y $]-\infty, 0[$. Como $\lim_{x\to 0} f(x) = 0 = f(0)$, f también es continua en 0, de modo que f es continua en todo punto. Por último, f no es derivable en 0, ya que si lo fuera,

$$\lim_{x\to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x\to 0} \frac{f(x)}{x}$$

existiría. Pero para $x > 0$, $f(x)/x$ es $+1$ y para $x < 0$ es -1, por lo que el límite no puede existir. ♦

6.3.4 Ejemplo *¿Debe ser continua la derivada de una función?*

Solución La respuesta es no, pero no es obvio encontrar un ejemplo; uno de ellos es la función

$$f(x) = x^2 \text{sen } \frac{1}{x} \quad \text{si } x \neq 0,$$

y

$$f(x) = 0 \quad \text{si } x = 0.$$

(Véase la figura 6.3-2).

Para demostrar la derivabilidad en cero, mostraremos que

$$\frac{f(x)}{x} \to 0 \quad \text{cuando } x \to 0.$$

Tenemos que $|f(x)/x| = |x \text{ sen}(1/x)| \leq |x| \to 0$ cuando $x \to 0$. Así, $f'(0)$ existe y es igual a cero. Por lo tanto, f es derivable en 0. Por las reglas de derivación del cálculo de una variable,

$$f'(x) = 2x \text{sen} \frac{1}{x} - \cos \frac{1}{x} \quad \text{si } x \neq 0.$$

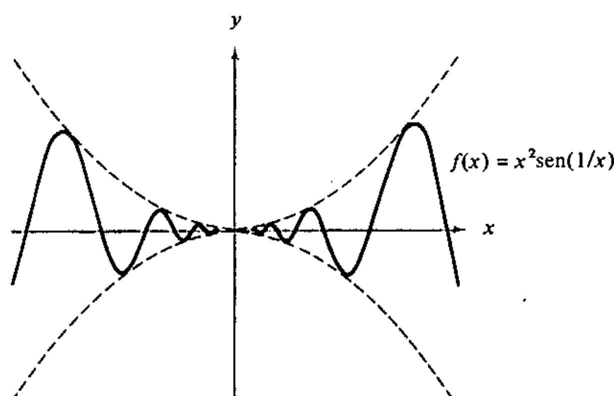


FIGURA 6.3-2 Esta función es derivable, pero la derivada es discontinua en 0

Cuando $x \to 0$, el primer término tiende a 0 pero el segundo oscila entre $+1$ y -1, de modo que $\lim_{x\to 0} f'(x)$ no existe. Así, f' existe pero no es continua. ♦

Suplemento de §6.3

Karl Weierstrass dio un ejemplo de una función continua $f(x)$ de una variable que no tiene derivada en ningún punto (véase el ejercicio 20, capítulo 5). Intuitivamente, esta función superpone un número infinito de picos, como los del ejemplo **6.3.3**. Uno podría preguntarse si tales funciones surgen en un contexto "práctico". Para nuestra sorpresa, la respuesta es sí. Un lugar donde las curvas "no derivables en ningún punto" son importantes es en el estudio del ***movimiento browniano***, el movimiento errático de pequeñas partículas suspendidas en un fluido, como el agua o el aire. El matemático estadounidense Norbert Wiener desarrolló un modelo eficaz para el estudio de este movimiento introduciendo una integración sobre un "espacio" de trayectorias. En este método, la "mayoría" de las trayectorias continuas son no derivables en ninguna parte y no tienen vectores tangentes.

Otro contexto en el que aparecen las curvas no derivables en ninguna parte es en el estudio de los fractales y los sistemas dinámicos. Por ejemplo, fijemos un número complejo c y sea $f_c : \mathbb{R}^2 \approx \mathbb{C} \to \mathbb{C}$ dada por $f_c(z) = z^2 + c$. Sea $J_c^\infty = \{z \mid f_c^n(z) \to \infty$ cuando $n \to \infty\}$, donde f_c^n indica la composición de f_c consigo misma n veces y sea $J_c = \partial(J_c^\infty)$. Para $c = 0$, está claro que J_c es un círculo, pero para $c \neq 0$ (aunque suficientemente cercano a 0) se sabe que J_c es ¡una curva continua pero no derivable en ninguna parte! J_c es el ***conjunto de Julia*** de f_c, que aparece en la figura 6.3-3.

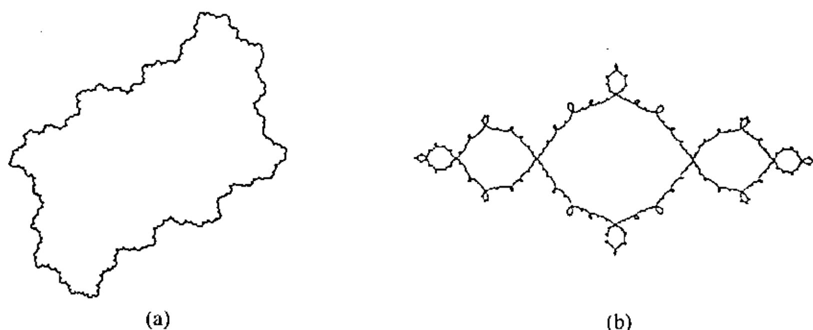


FIGURA 6.3-3 (a) El conjunto de Julia de $f(z) = z^2 + \frac{1}{2}i$, que es una curva simple cerrada pero que no es derivable en ningún punto; (b) el conjunto de Julia de $f(z) = z^2 - 1$, que contiene infinitas curvas cerradas

El conjunto de los valores c tales que J_c es arco-conexo se llama ***conjunto de Mandelbrot***, el cual se muestra en la figura 6.3-4 (el conjunto de Mandelbrot es también el conjunto de valores c para los que cero no tiende a infinito bajo la iteración de f_c; sin embargo, esto no es evidente). Un profundo teorema de Douady y Hubbard postula que *el conjunto de Mandelbrot es conexo*. Éste es un ejemplo de un conjunto muy complicado cuya conexión está lejos de ser evidente; un acercamiento a este conjunto mediante el computador muestra lo sutil que es. Para mayor información, bibliografía y razones de la importancia práctica de este tema, véase R. Devaney, *An Introduction to Chaotic Dynamical Systems* (Addison-Wesley, Reading, Mass., 1985) y la bibliografía incluida en dicho texto.

Ejercicios de §6.3

1. Sea $f(x) = x^2$ si x es irracional y $f(x) = 0$ si x es racional. ¿Es f continua en 0? ¿Es derivable en 0?

2. ¿Es la condición local de Lipschitz del teorema **6.3.1** suficiente para garantizar la diferenciabilidad?

3. ¿Debe existir la derivada de una función continua en su máximo?

4. Sea $f(x) = x\,\text{sen}(1/x), x \neq 0$ y $f(0) = 0$. Investíguese la continuidad y derivabilidad de f en 0.

5. Encuéntrese el vector tangente a la curva $c(t) = (3t^2, e^t, t + t^2)$ en el punto correspondiente a $t = 1$.

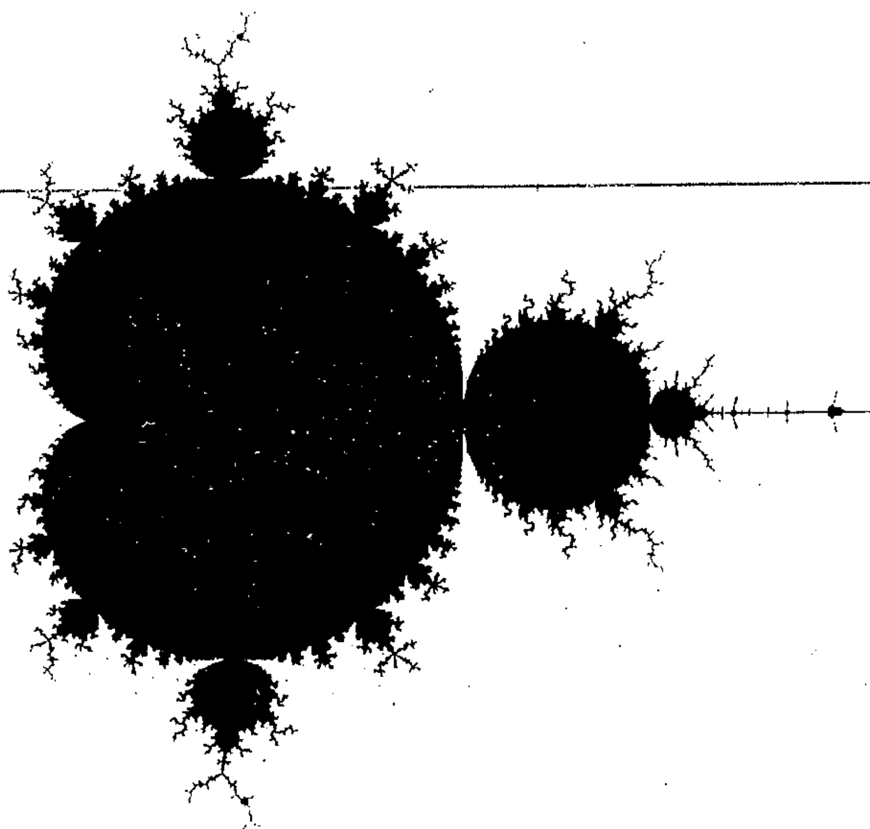


FIGURA 6.3-4 El conjunto de Mandelbrot (reimpreso con autorización de B.B. Mandelbrot, *The Fractal Geometry of Nature*, W.H. Freeman, Nueva York, 1983)

§6.4 Condiciones para la diferenciabilidad

Puesto que la matriz jacobiana proporciona una herramienta de cálculo eficaz, sería útil saber si la existencia de las derivadas parciales usuales implica que la diferencial $\mathbf{D}f$ existe. Esto, por desgracia, no es cierto en general. Por ejemplo, sea $f: \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x, y) = x$ si $y = 0$, $f(x, y) = y$ si $x = 0$ y $f(x, y) = 1$ en los demás casos. Entonces $\partial f/\partial x$ y $\partial f/\partial y$ existen en (0, 0) y son iguales a 1. Sin embargo, f no es continua en (0, 0) (¿por qué?), y entonces la diferencial $\mathbf{D}f$ no puede ni siquiera existir en (0, 0). Véase la figura 6.4-1 (véanse los ejemplos y ejercicios para otros ejemplos más exóticos).

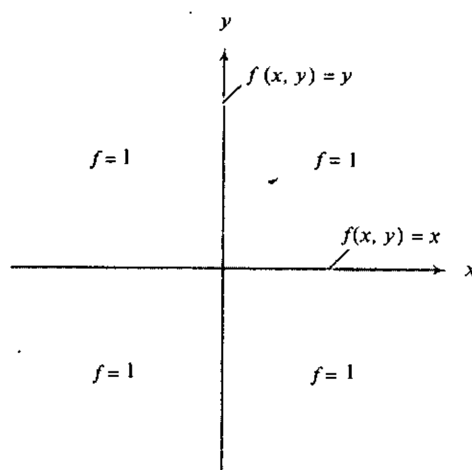


FIGURA 6.4-1 Las derivadas parciales de esta función existen en (0, 0), pero la función no es continua en (0, 0)

Es bastante sencillo comprender tal comportamiento. Las derivadas parciales sólo dependen de lo que ocurre en la dirección de los ejes x e y, mientras que la definición de $\mathbf{D}f$ implica el comportamiento combinado de f en toda una vecindad de un punto dado. Sin embargo, podemos afirmar lo siguiente.

6.4.1 Teorema *Sean $A \subset \mathbb{R}^n$ un conjunto abierto y $f: A \subset \mathbb{R}^n \to \mathbb{R}^m$. Supóngase que $f = (f_1, \ldots, f_m)$. Si cada derivada parcial $\partial f_j/\partial x_i$ existe y es continua en A, entonces f es diferenciable en A.*

Las derivadas parciales de una función miden su variación en las direcciones particulares paralelas a los ejes. Las ***derivadas direccionales*** hacen lo mismo en otras direcciones.

6.4.2 Definición *Sea f una función escalar definida en una vecindad de $x_0 \in \mathbb{R}^n$ y sea $e \in \mathbb{R}^n$ un vector unitario. Entonces*

$$\frac{d}{dt} f(x_0 + te)|_{t=0} = \lim_{t \to 0} \frac{f(x_0 + te) - f(x_0)}{t}$$

*es la **derivada direccional** de f en x_0 en la dirección e.*

Por esta definición, vemos que la derivada direccional es la variación de f en la dirección e, y, para $n = 2$, da la pendiente de la gráfica de f en esa dirección; véase la figura 6.4-2.

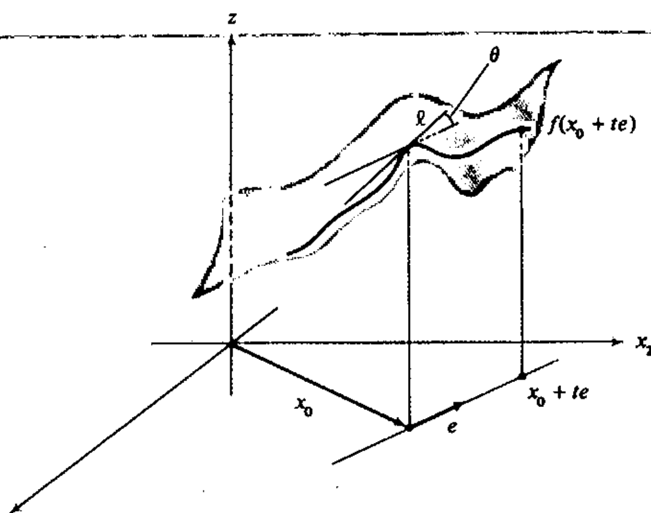


FIGURA 6.4-2 Pendiente de l = $\tan\theta$ = derivada direccional

Afirmamos que *la derivada direccional en la dirección de e es igual a* $\mathbf{D}f(x_0) \cdot e$. Para ver esto, considérese la definición de $\mathbf{D}f(x_0)$ con $x = x_0 + te$: dado $\varepsilon > 0$,

$$\left\| \frac{f(x_0 + te) - f(x_0)}{t} - \mathbf{D}f(x_0) \cdot e \right\| \leq \varepsilon \|e\|$$

si $|t|$ es suficientemente pequeño. Así, si f es diferenciable en x_0, entonces las derivadas direccionales también existen y están dadas por

$$\lim_{t\to 0} \frac{f(x_0 + te) - f(x_0)}{t} = \mathbf{D}f(x_0) \cdot e.$$

En particular, obsérvese que $\partial f/\partial x_i$ es la derivada de f en la dirección del i-ésimo eje de coordenadas (con $e = e_i = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$).

Para una función $f\colon \mathbb{R}^2 \to \mathbb{R}$, las derivadas direccionales $\mathbf{D}f(x_0) \cdot e$ se pueden usar para determinar el plano tangente a la gráfica de f (compárese la figura 6.1-2). Es decir, la recta l dada por $z = f(x_0) + \mathbf{D}f(x_0) \cdot te$ es tangente a la gráfica de f, pues, como en la figura 6.4-2, $\mathbf{D}f(x_0) \cdot e$ es la variación de f en la dirección e. Así, el plano tangente a la gráfica de f en $(x_0, f(x_0))$ se puede describir mediante la ecuación

$$z = f(x_0) + \mathbf{D}f(x_0) \cdot (x - x_0)$$

(véase la figura 6.4-3). Puesto que no hemos definido el concepto de plano tangente a una superficie, adoptaremos esta ecuación como *definición* del plano tangente.

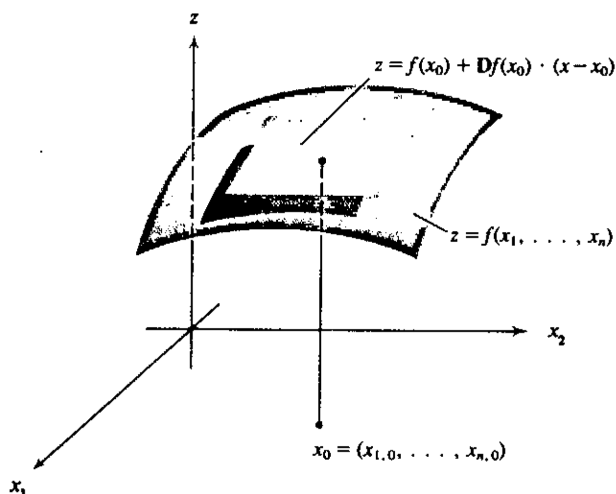


FIGURA 6.4-3 El plano tangente a la gráfica de una función

6.4.3 Ejemplo *Muéstrese que la existencia de todas las derivadas direccionales en un punto no implica necesariamente la diferenciabilidad.*

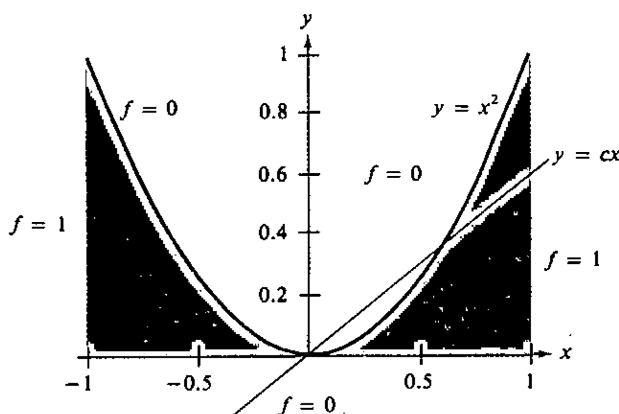


FIGURA 6.4-4 La existencia de las derivadas direccionales no implica la diferenciabilidad, ni siquiera la continuidad

Solución Definamos una función $f: \mathbb{R}^2 \to \mathbb{R}$ como $f(x, y) = 1$ si $0 < y < x^2$ y como $f(x, y) = 0$ en los demás casos. Entonces f se anula a lo largo de los ejes horizontal y vertical. Cualquier otra línea que pase por el origen permanece en la región donde $f = 0$ en una corta distancia a ambos lados del origen. Así, f es constante e igual a cero en algún intervalo a ambos lados del origen a lo largo de tal recta. Esto muestra que la derivada direccional en el origen en la dirección de esa recta es 0. Todas las derivadas direccionales existen y se anulan en el origen, pero la función ni siquiera es continua, mucho menos diferenciable, en el origen. Véase la figura 6.4-4.

Un ejemplo un poco más complicado pero similar es la función $f: \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x,y) = \frac{xy}{x^2 + y} \quad \text{si} \quad x^2 \neq -y$$

y

$$f(x,y) = 0 \quad \text{si} \quad x^2 = -y.$$

Si $e = (e_1, e_2)$, donde $e_2 \neq 0$,

$$\frac{1}{t} f(te_1, te_2) = \frac{1}{t} \frac{t^2 e_1 e_2}{t^2 e_1^2 + te_2} = \frac{e_1 e_2}{te_1^2 + e_2} \to e_1$$

cuando $t \to 0$; el caso $e_2 = 0$ da como resultado cero. Así, cada derivada direccional existe en (0, 0), pero f no es continua en (0, 0), ya que para x^2 cerca de $-y$ y x, y pequeños, f es muy grande (dejamos los detalles al lector). ◆

Este ejemplo muestra que la existencia de todas las derivadas direccionales *no* es una definición conveniente de diferenciabilidad, ya que ni siquiera implica continui-

dad. Ésta es la razón por la que se adopta un concepto más restringido en la definición **6.1.1**.

6.4.4 Ejemplo *Sea $f(x, y) = x^2 + y$. Calcúlese la ecuación del plano tangente a la gráfica de f en $x = 1$, $y = 2$.*

Solución En este caso $\mathbf{D}f(x, y)$ tiene la matriz

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = (2x, 1),$$

y entonces $\mathbf{D}f(1, 2) = (2, 1)$. Así, la ecuación del plano tangente es

$$z = 3 + (2, 1)\begin{pmatrix} x-1 \\ y-2 \end{pmatrix} = 3 + 2(x-1) + (y-2);$$

es decir,

$$2x + y - z = 1. \quad \blacklozenge$$

Ejercicios de §6.4

1. Úsese el teorema **6.4.1** para mostrar que $f(x, y)$ dada por

$$f(x,y) = \frac{(xy)^2}{\sqrt{x^2+y^2}}, \quad (x,y) \neq (0,0)$$

y

$$f(x,y) = 0, \quad (x,y) = (0,0)$$

es diferenciable en $(0, 0)$.

2. Analícese la diferenciabilidad de

$$f(x,y) = \frac{xy}{\sqrt{x^2+y^2}}$$

en $(0, 0)$ si $f(0, 0) = 0$.

3. Encuéntrese el plano tangente a la gráfica de $z = x^2 + y^2$ en $(0, 0)$.

4. Encuéntrese la ecuación del plano tangente a la gráfica de $z = x^3 + y^4$ en $x = 1$, $y = 3$.

5. Encuéntrese una función $f: \mathbb{R}^2 \to \mathbb{R}$ que sea diferenciable en cada punto pero cuyas derivadas parciales no sean continuas en $(0, 0)$.

§6.5 La regla de la cadena

Como vimos en **§4.7**, una de las técnicas de derivación más importantes es la regla de la cadena (regla de la "función de una función"). Por ejemplo, para derivar $(x^3 + 3)^6$, sea $y = x^3 + 3$; primero derivamos y^6, obteniendo $6y^5$, y después multiplicamos por la derivada de $x^3 + 3$ para obtener la respuesta final $6(x^3 + 3)^5 3x^2$. Existe un proceso análogo para funciones de varias variables. Por ejemplo, si u, v y f son funciones escalares de dos variables, entonces

$$\frac{\partial}{\partial x} f(u(x,y), v(x,y)) = \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x}.$$

El siguiente teorema incluye todos los casos anteriores como casos particulares.

6.5.1 Regla de la cadena *Sean $A \subset \mathbb{R}^n$ abierto y $f: A \to \mathbb{R}^m$ diferenciable en $x_0 \in A$. Sean $B \subset \mathbb{R}^n$ abierto, $f(A) \subset B$ y $g: B \to \mathbb{R}^p$ diferenciable en $f(x_0)$. Entonces la composición $g \circ f$ es diferenciable en x_0 y $\mathbf{D}(g \circ f)(x_0) = \mathbf{D}g(f(x_0)) \circ \mathbf{D}f(x_0)$.*

Obsérvese que el miembro derecho de la última igualdad está bien definido, pues $\mathbf{D}f(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ y $\mathbf{D}g(f(x_0)) : \mathbb{R}^m \to \mathbb{R}^p$ y, por lo tanto, su composición como transformaciones lineales está definida. La regla de la cadena se llama también ***teorema de la composición de funciones***, pues nos dice cómo derivar una composición de funciones.

Como el producto de dos matrices corresponde a la composición de las transformaciones lineales correspondientes que representan, la regla de la cadena se puede reescribir diciendo que *la matriz jacobiana de $g \circ f$ en $x = (x_1, \ldots, x_n)$ es el producto de la matriz jacobiana de g evaluada en $f(x)$ con la matriz jacobiana de f evaluada en x (en ese orden)*. Así, si $h = g \circ f$ e $y = f(x)$, entonces

$$\mathbf{D}h(x) = \begin{pmatrix} \dfrac{\partial g_1}{\partial y_1} & \cdots & \dfrac{\partial g_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_p}{\partial y_1} & \cdots & \dfrac{\partial g_p}{\partial y_m} \end{pmatrix} \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix}$$

donde las $\partial g_i/\partial y_j$ se evalúan en $y = f(x)$ y las $\partial f_i/\partial x_j$ en x. Si desarrollamos esto, obtenemos, por ejemplo,

$$\frac{\partial h_1}{\partial x_1} = \sum_{j=0}^{m} \frac{\partial g_1}{\partial y_j}\,\frac{\partial f_j}{\partial x_1}.$$

Esta situación aparece cuando "cambiamos de variable". Por ejemplo, supongamos que $f(x, y)$ es una función escalar y $x = r\cos\theta$, $y = r\,\mathrm{sen}\,\theta$ definen las nuevas variables r, θ en coordenadas polares. Formamos la función

$$h(r, \theta) = f(r\cos\theta, r\,\mathrm{sen}\,\theta).$$

Entonces

$$\frac{\partial h}{\partial r} = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\mathrm{sen}\,\theta$$

y

$$\frac{\partial h}{\partial \theta} = -\frac{\partial f}{\partial x} r\,\mathrm{sen}\,\theta + \frac{\partial f}{\partial y} r\cos\theta.$$

El lector debe obtener fórmulas similares para las coordenadas cilíndricas (r, φ, θ), donde $x = r\cos\theta\,\mathrm{sen}\,\varphi$, $y = r\,\mathrm{sen}\,\theta\,\mathrm{sen}\,\varphi$, $z = r\cos\varphi$ (analizaremos las coordenadas esféricas más adelante, en **§9.5**).

Otro ejemplo puede ayudar a aclarar las cosas. Supongamos que tenemos las funciones $u(x, y)$, $v(x, y)$, $w(x, y)$ y $f(u, v, w)$ y que formamos la función $h(x, y) = f(u(x, y), v(x, y), w(x, y))$. La regla de la cadena nos da

$$\frac{\partial h}{\partial x} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial x}.$$

La idea detrás de la demostración de esta fórmula (como ejemplo ilustrativo) es la siguiente. Escribimos

$$\begin{aligned}\frac{h(x+\Delta x, y) - h(x, y)}{\Delta x} &= \frac{f(u(x+\Delta x, y), v(x+\Delta x, y), w(x+\Delta x, y))}{\Delta x}\\ &\quad - \frac{f(u(x, y), v(x+\Delta x, y), w(x+\Delta x, y))}{\Delta x}\\ &\quad + \frac{f(u(x, y), v(x+\Delta x, y), w(x+\Delta x, y))}{\Delta x}\\ &\quad - \frac{f(u(x, y), v(x, y), w(x+\Delta x, y))}{\Delta x}\\ &\quad + \frac{f(u(x, y), v(x, y), w(x+\Delta x, y))}{\Delta x}\\ &\quad - \frac{f(u(x, y), v(x, y), w(x, y))}{\Delta x}.\end{aligned}$$

Usamos $f(u + \Delta u, v, w) - f(u, v, w) \approx \Delta u \partial f/\partial u$ y vemos que la expresión anterior es, aproximadamente,

$$\frac{\partial f}{\partial u}\frac{\Delta u}{\Delta x} + \frac{\partial f}{\partial v}\frac{\Delta v}{\Delta x} + \frac{\partial f}{\partial w}\frac{\Delta w}{\Delta x}.$$

Así, haciendo $\Delta x \to 0$ se obtiene la fórmula. ¿Puede el lector determinar cuáles son los detalles del argumento que deben aclararse para que la demostración esté completa?

6.5.2 Ejemplo *Verifíquese la regla de la cadena para* $f(u, v, w) = u^2v + wv^2$ *y* $u = xy$, v sen x, $w = e^x$.

Solución En este caso, $h(x, y) = f(u(x, y), v(x, y), w(x, y))$ está dada por

$$h(x, y) = x^2y^2 \operatorname{sen} x + e^x \operatorname{sen}^2 x,$$

y entonces, directamente,

$$\frac{\partial h}{\partial x} = 2xy^2 \operatorname{sen} x + x^2y^2 \cos x + e^x \operatorname{sen}^2 x + e^x \, 2 \operatorname{sen} x \cos x.$$

Por otro lado, la regla de la cadena implica

$$\frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial x} = 2uv\frac{\partial u}{\partial x} + u^2\frac{\partial v}{\partial x} + 2wv\frac{\partial v}{\partial x} + v^2\frac{\partial w}{\partial x}$$

$$= 2xy^2 \operatorname{sen} x + x^2y^2 \cos x + 2e^x \operatorname{sen} x \cos x + e^x \operatorname{sen}^2 x,$$

que es el mismo resultado. La fórmula para $\partial h/\partial y$ se puede verificar de modo análogo. ♦

6.5.3 Ejemplo *Sean* $f: \mathbb{R} \to \mathbb{R}$ *y* $F: \mathbb{R}^2 \to \mathbb{R}$ *dadas por* $F(x, y) = f(xy)$. *Verifíquese que*

$$x\frac{\partial F}{\partial x} = y\frac{\partial F}{\partial y}.$$

Solución Por la regla de la cadena,

$$\frac{\partial F}{\partial x} = f'(xy)y \qquad \text{y} \qquad \frac{\partial F}{\partial y} = f'(xy)x,$$

por lo que la afirmación es clara. ♦

6.5.4 Ejemplo *Supóngase que* $\gamma : [0, 1] \to \mathbb{R}^n$ *es una parametrización diferenciable de una curva en* $\mathbb{R}^n$ *y que* f *es una función diferenciable de* $\mathbb{R}^n$ *en* $\mathbb{R}$. *Sea* $h(t) = f(\gamma(t))$. *Muéstrese que* $h'(t) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle$ *y explíquese cómo esto generaliza la fórmula obtenida anteriormente para derivadas direccionales (véase la figura 6.5-1).*

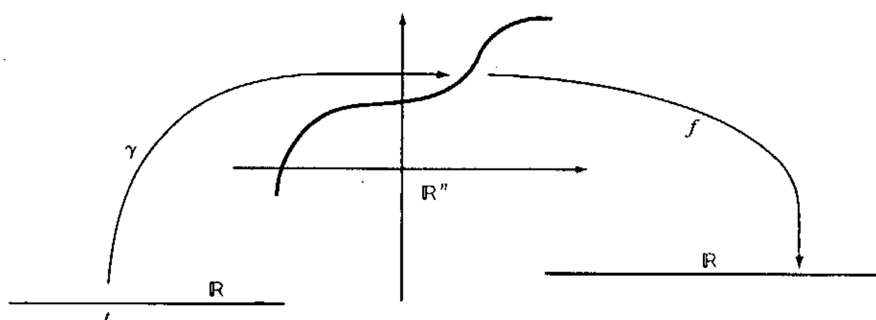


FIGURA 6.5-1 Una función definida a lo largo de una curva parametrizada

Solución Como $h(t) = (f \circ \gamma)(t)$, la regla de la cadena implica

$$h'(t) = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right] \begin{bmatrix} x_1'(t) \\ x_2'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}$$

$$= \frac{\partial f}{\partial x_1} x_1'(t) + \frac{\partial f}{\partial x_2} x_2'(t) + \cdots + \frac{\partial f}{\partial x_n} x_n'(t).$$

Todas las derivadas parciales se evalúan en el punto $\gamma(t) = (x_1(t), x_2(t), \ldots, x_n(t))$. Esto es exactamente igual al producto escalar del gradiente de f en ese punto con el vector velocidad $\gamma'(t)$ en ese punto. La derivada direccional de f en el punto v_0 en la dirección del vector unitario $e = (a_1, a_2, \ldots, a_n)$ es el caso particular en que la curva es la línea recta $\gamma(t) = v_0 + te$. En ese caso, $x_k(t) = \gamma_k(t) = v_{0k} + ta_k$ y podemos recuperar la misma fórmula anterior:

$$\langle \nabla f(\gamma(0)), \gamma'(0) \rangle = \langle \nabla f(v_0), e \rangle$$

pues $\gamma(0) = v_0$ y $\gamma'(t) = e$ para todo t. ♦

Ejercicios de §6.5

1. Desarróllese la regla de la cadena para

$$h(x, y, z) = f(u(x, y, z),\ v(x, y), w(y, z)).$$

2. Verifíquese la regla de la cadena para

$$u(x,y,z) = xe^y, \quad v(x,y,z) = (\operatorname{sen} x)yz$$

y

$$f(u,v) = u^2 + v\operatorname{sen} u$$

con

$$h(x,y,z) = f(u(x,y,z), v(x,y,z)).$$

3. Sea $F(x, y) = f(x^2 + y^2)$. Muéstrese que $x(\partial F/\partial y) = y(\partial F/\partial x)$.

4. Desarróllese la regla de la cadena que relaciona las coordenadas rectangulares con las coordenadas esféricas en tres dimensiones.

5. Sean $f: \mathbb{R} \to \mathbb{R}$ y $F: \mathbb{R}^2 \to \mathbb{R}$ diferenciables tales que $F(x, f(x)) = 0$ y $\partial F/\partial y \neq 0$. Demuéstrese que $f'(x) = -\dfrac{\partial F/\partial x}{\partial F/\partial y}$, donde $y = f(x)$.

§6.6 Regla del producto y gradientes

Otra regla bien conocida del cálculo diferencial es la ***regla del producto***, o ***regla de Leibniz***.

6.6.1 Proposición *Sean $A \subset \mathbb{R}^n$ abierto y $f: A \to \mathbb{R}^m$ y $g: A \to \mathbb{R}$ funciones diferenciables. Entonces gf es diferenciable, y para $x \in A$, $\mathbf{D}(gf)(x): \mathbb{R}^n \to \mathbb{R}^m$ está dada por $\mathbf{D}(gf)(x) \cdot e = g(x)(\mathbf{D}f(x) \cdot e) + (\mathbf{D}g(x) \cdot e)f(x)$ para todo $e \in \mathbb{R}^n$ (esto tiene sentido, pues $g(x) \in \mathbb{R}$ y $\mathbf{D}g(x) \cdot e \in \mathbb{R}$).*

A veces abreviamos este resultado diciendo que

$$\mathbf{D}(gf) = g\mathbf{D}f + (\mathbf{D}g)f,$$

pero el significado preciso es como en el enunciado en la proposición. El lector ya está familiarizado con la regla del producto por **§4.7** y por el cálculo elemental. En términos de componentes, la proposición es, de hecho, simplemente la regla del producto usual

$$\frac{\partial}{\partial x_i}(gf_k) = g\frac{\partial f_k}{\partial x_i} + \frac{\partial g}{\partial x_i}f_k.$$

Para los cocientes, tenemos un resultado análogo. Si $g \neq 0$, entonces

$$\mathbf{D}\left(\frac{f}{g}\right) = \frac{g\mathbf{D}f - (\mathbf{D}g)f}{g^2}.$$

Para demostrar esta fórmula, basta, por la proposición **6.6.1**, demostrarla para el caso $1/g$. Esto se reduce a un problema de cálculo elemental, que ya hemos pedido al lector que resolviera en **§4.7.**

Otras reglas de derivación se agrupan en la afirmación de que **D** es *lineal*; es decir, $\mathbf{D}(f+g) = \mathbf{D}f + \mathbf{D}g$ y $\mathbf{D}(\lambda f) = \lambda\mathbf{D}f$ para $\lambda \in \mathbb{R}$ constante.

Consideremos la *geometría de los gradientes* con más detalle. Sea $f: A \subset \mathbb{R}^n \to \mathbb{R}$ diferenciable. El gradiente es

$$\operatorname{grad} f(x) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right).$$

Por lo tanto, la derivada direccional en la dirección h es el producto escalar

$$\begin{aligned}\mathbf{D}f(x) \cdot h &= \langle \operatorname{grad} f(x), h\rangle \\ &= \text{variación de } f \text{ en el punto } x \text{ en la dirección } h.\end{aligned}$$

Considérese una "superficie" S definida por $f(x)$ = constante. *Ahora mostraremos que* grad $f(x)$ *es ortogonal a esta superficie* (nuestro argumento se basará en la naturaleza precisa de esta superficie, que analizaremos en **§7.7**). Para demostrar esto, considérese una curva $c(t)$ en S con vector tangente $c'(0)$, donde $c(0) = x_0$. Afirmamos que

$$\langle \operatorname{grad} f(x_0), c'(0)\rangle = 0.$$

Como $c(t) \in S, f(c(t))$ = constante en t. Derivamos esta expresión usando la regla de la cadena para obtener

$$\mathbf{D}f(c(t)) \cdot c'(t) = 0.$$

Hacemos $t = 0$ y usamos que $\mathbf{D}f(x) \cdot h = \langle \operatorname{grad} f(x), h\rangle$, para obtener la relación deseada. Véase la figura 6.6-1.

Podemos describir el plano tangente a la superficie S definida por $f(x)$ = constante en un punto x_0 en S como $\langle \operatorname{grad} f(x_0), x - x_0\rangle$, pues grad $f(x_0)$ es ortogonal a S.

También es evidente de la ecuación

$$\langle \operatorname{grad} f(x_0), h\rangle = \|\operatorname{grad} f(x_0)\| \cos\theta$$

(donde $\|h\| = 1$ y θ es el ángulo entre grad $f(x_0)$ y h) que el vector $f(x_0)$ apunta en la dirección donde f cambia más rápidamente. Esto no es absurdo, pues si suponemos que f representa la función altura de una montaña, entonces f = constante es un contorno de

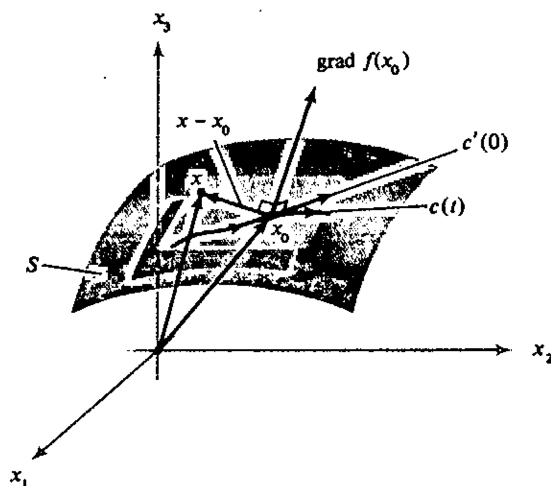


FIGURA 6.6-1 El gradiente es ortogonal a los conjuntos de nivel

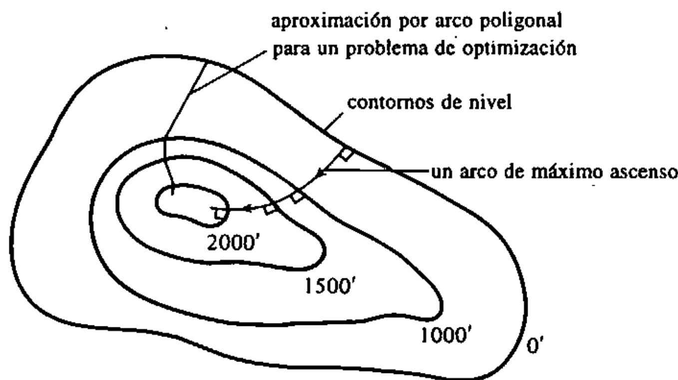


FIGURA 6.6-2 La dirección de máximo ascenso es ortogonal a los contornos de nivel

nivel. Para subir o descender de la montaña lo más pronto posible, debemos caminar en forma perpendicular a los contornos de nivel (figura 6.6-2).

Estos hechos pueden ser valiosos en problemas de optimización, donde se tiene una función $f(x_1, \ldots, x_n)$ y el problema es maximizar u "optimizar" f mediante algún esquema práctico. Un método común es el de tomar un punto de prueba x_0 y continuar a lo largo de una línea recta en la dirección del gradiente de f para alcanzar un nuevo punto en el que f sea mayor (al menos si no vamos demasiado lejos), y repetir el procedimiento.

6.6.2 Ejemplo *Determínese la normal a la superficie* $x^2 + y^2 + z^2 = 3$ *en* (1, 1, 1).

Solución En este caso, $f(x, y, z) = x^2 + y^2 + z^2$ tiene gradiente $f = (2x, 2y, 2z)$ que vale (2, 2, 2) en (1, 1, 1). Al normalizar, un vector normal unitario es $(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$. ♦

6.6.3 Ejemplo *Determínese la dirección de máxima variación de* $f(x, y, z) = x^2y$ sen z *en* (3, 2, 0).

Solución La dirección pedida es la del vector gradiente, que es ($2xy$ sen z, x^2 sen z, x^2y cos z) o (0, 0, 18) en (3, 2, 0). Al normalizar obtenemos el vector director unitario (0, 0, 1). ♦

6.6.4 Ejemplo *¿Cuál es el plano tangente a la superficie* $x^2 - y^2 + xz = 2$ *en* (1, 0, 1)*?*

Solución En este caso $f(1, 0, 1) = (3, 0, 1)$, de modo que el plano tangente tiene la ecuación $\langle(x - 1, y, z - 1), (3, 0, 1)\rangle = 0$; es decir, $3x + z = 4$. ♦

Ejercicios de §6.6

1. Demuéstrese que

$$\frac{d}{dt} f(x_0 + th)\bigg|_{t=0} = \mathbf{D}f(x_0) \cdot h$$

mediante la regla de la cadena, donde $f: \mathbb{R}^m \to \mathbb{R}^n$.

2. Determínese un vector normal unitario a la superficie $x^2 - y^2 + xyz = 1$ en (1, 0, 1).

3. Determínese la ecuación del plano tangente a la superficie $x^2 - y^2 + xyz = 1$ en (1, 0, 1).

4. ¿En qué dirección crece más rápidamente $f(x, y) = e^{x^2}$?

5. Sean $f: \mathbb{R}^n \to \mathbb{R}$, $g: \mathbb{R}^n \to \mathbb{R}$. Muéstrese que grad $(fg) = f$ grad $g + g$ grad f.

6. Muéstrese que grad f, como normal al plano tangente, es una descripción más general de dicho plano que la descripción dada en **§6.4.**

§6.7 El teorema del valor medio

Dos teoremas importantes del cálculo son el teorema del valor medio y el teorema de Taylor. Primero, analizaremos el teorema del valor medio. En **§4.7** recordamos la demostración del teorema del valor medio del cálculo elemental, el cual establece que si $f: [a, b] \to \mathbb{R}$ es continua y si f es derivable en $]a, b[$, existe un punto $c \in]a, b[$ tal que $f(b) - f(a) = f'(c)(b - a)$, donde $f' = df/dx$. Por desgracia, para $f: A \subset \mathbb{R}^n \to \mathbb{R}^m$, esta versión del teorema del valor medio, interpretada ingenuamente, no es cierta. Por ejemplo, considérese $f: \mathbb{R} \to \mathbb{R}^2$ dada por $f(x) = (x^2, x^3)$. Intentemos buscar c tal que $0 \le c \le 1$ y $f(1) - f(0) = \mathbf{D}f(c)(1 - 0)$. Esto significa que $(1, 1) - (0, 0) = (2c, 3c^2)$ y por lo tanto $2c = 1$ y $3c^2 = 1$. Está claro que no existe c que satisfaga estas ecuaciones.

La experiencia nos lleva a pensar que alguna restricción produciría un teorema válido. De hecho, la versión anterior será válida si *f es escalar*. Para enunciar el teorema correcto, primero precisaremos el significado de que "c está entre x e y" para $c, x, y \in \mathbb{R}^n$. Decimos que c está ***en el segmento de recta que une** x **con** y*, o ***está entre** x **e** y*, si $c = (1 - \lambda)x + \lambda y$ para algún $0 \le \lambda \le 1$, como en la figura 6.7-1.

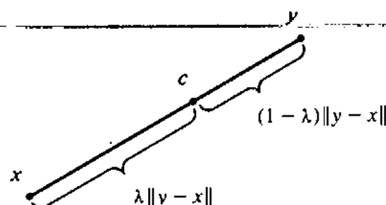


FIGURA 6.7-1 El punto c está en el segmento que une x e y

6.7.1 Teorema del valor medio

i. *Supóngase que $f: A \subset \mathbb{R}^n \to \mathbb{R}$ es diferenciable en un conjunto abierto A. Para cualesquiera $x, y \in A$ tales que el segmento de recta que une x con y está en A (lo que no ocurre necesariamente para todo x, y), existe un punto c en dicho segmento tal que*

$$f(y) - f(x) = \mathbf{D}f(c) \cdot (y - x).$$

ii. *Supóngase que $f: A \subset \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en el conjunto abierto A. Supóngase que el segmento de recta que une x con y está contenido en A y que $f = (f_1, \ldots, f_m)$. Entonces existen puntos $c_1, \ldots, c_m$ en dicho segmento tales que*

$$f_i(y) - f_i(x) = \mathbf{D}f_i(c_i)(y - x), \quad i = 1, \ldots, m.$$

Una formulación alternativa del teorema del valor medio está dada en el ejemplo resuelto **6.5** al final del capítulo.

6.7.2 Ejemplo *Un conjunto $A \subset \mathbb{R}^n$ es **convexo** si para cada $x, y \in A$, el segmento que une x con y está contenido en A. Véase la figura 6.7-2. Sea $A \subset \mathbb{R}^n$ un conjunto abierto convexo y $f: A \to \mathbb{R}^m$ diferenciable. Si $\mathbf{D}f = 0$ en A, muéstrese que f es constante en A (las generalizaciones de este resultado aparecen en el ejercicio 9 al final del capítulo).*

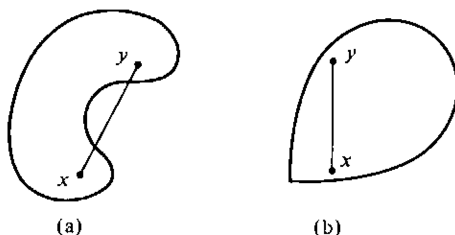


FIGURA 6.7-2 (a) No convexo; (b) convexo

Solución Si $x, y \in A$, entonces para cada componente f_i existe un vector c_i tal que

$$f_i(y) - f_i(x) = \mathbf{D}f_i(c_i)(y - x).$$

Como $\mathbf{D}f = 0$, $\mathbf{D}f_i = 0$ para cada i (¿por qué?), de modo que $f_i(y) = f_i(x)$. Esto implica que $f(y) = f(x)$, lo que significa que f es constante. ♦

6.7.3 Ejemplo *Supóngase que $f: [0, \infty[\to \mathbb{R}$ es continua, $f(0) = 0$, f es derivable en $]0, \infty[$ y f' es no decreciente. Demuéstrese que $g(x) = f(x)/x$ es no decreciente para $x > 0$.*

Solución Por §**4.7**, o directamente del teorema del valor medio, vemos que una función $h: \mathbb{R} \to \mathbb{R}$ es no decreciente si $h'(x) \geq 0$, pues $x \leq y$ implica que

$$h(y) - h(x) = h'(c)(y - x) \geq 0.$$

Ahora

$$g'(x) = \frac{xf'(x) - f(x)}{x^2}$$

y existe c entre 0 y x tal que

$$f(x) = f(x) - f(0) = f'(c) \cdot x \leq xf'(x),$$

pues $0 < c < x$ implica que $f'(x) \geq f'(c)$. Así, $xf'(x) - f'(x) \geq 0$, de modo que $g' \geq 0$, lo que implica que g es no decreciente. ♦

Ejercicios de §6.7

1. Si $f: \mathbb{R} \to \mathbb{R}$ es derivable y tal que $f'(x) > 0$, demuéstrese que f es (estrictamente) creciente.

2. Demuéstrese la siguiente versión (débil) de la ***regla de l'Hôpital***: Si f' y g' existen en x_0, $g'(x_0) \neq 0$ y $f(x_0) = 0 = g(x_0)$, entonces $\lim_{x \to x_0}[f(x)/g(x)] = f'(x_0)/g'(x_0)$.

3. Utilícese el ejercicio 2 para evaluar

   a. $\lim_{x \to 0}[(\operatorname{sen} x)/x]$

   b. $\lim_{x \to 0}[(e^x - 1)/x]$

4. ¿Cuáles de los siguientes conjuntos son convexos?

   a. $\{(x, y) \in \mathbb{R}^2 \mid y \geq 0\}$

   b. $\{x \in \mathbb{R}^n \mid 0 < \|x\| < 1\}$

   c. $\mathbb{R} \setminus \{0\}$

5. Sean $f: A \subset \mathbb{R}^n \to \mathbb{R}$ diferenciable y A convexo y supóngase que $\|\text{grad } f(x)\| \leq M$ para $x \in A$. Demuéstrese que $|f(x) - f(y)| \leq M\|x - y\|$ para $x, y \in A$. ¿Será esto cierto si A no es convexo?

6. Sea $f: \mathbb{R} \to \mathbb{R}$ derivable. Supóngase que para todo $x \in \mathbb{R}$, $0 \leq f'(x) \leq f(x)$. Muéstrese que $g(x) = e^{-x} f(x)$ es decreciente. Si f se anula en algún punto, conclúyase que f es cero.

§6.8 Teorema de Taylor y derivadas de orden superior

A continuación analizamos la fórmula de Taylor para funciones $f: A \subset \mathbb{R}^n \to \mathbb{R}^m$; pero para ello, primero analizaremos las derivadas de orden superior. Para $f: \mathbb{R}^n \to \mathbb{R}$ no hay

problema en definir las derivadas parciales de orden superior; simplemente iteramos el proceso de derivación parcial

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial}{\partial x_1}\left(\frac{\partial}{\partial x_2} f\right),$$

etcétera. Sin embargo, con respecto a la diferencial como transformación lineal, necesitamos un poco más de cuidado.

La diferencial segunda se obtiene al diferenciar $\mathbf{D}f$, si existe, lo que hacemos del modo siguiente.

6.8.1 Definición *Sea $L(\mathbb{R}^n, \mathbb{R}^m)$ el espacio de transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$ (si elegimos una base en $\mathbb{R}^n$ y otra en $\mathbb{R}^m$, entonces $L(\mathbb{R}^n, \mathbb{R}^m)$ se puede identificar con las matrices $m \times n$ y por lo tanto con $\mathbb{R}^{nm}$). Ahora, $\mathbf{D}f : A \to L(\mathbb{R}^n, \mathbb{R}^m)$; es decir, en cada $x \in A$ obtenemos una transformación lineal $\mathbf{D}f(x)$. Si diferenciamos $\mathbf{D}f$ en x_0, obtenemos una transformación lineal de $\mathbb{R}^n$ en $L(\mathbb{R}^n, \mathbb{R}^m)$, por definición de la diferencial. Escribimos $\mathbf{D}(\mathbf{D}f)(x_0) = \mathbf{D}^2 f(x_0)$ y definimos la transformación $B_{x_0} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ como $B_{x_0}(x_1, x_2) = [\mathbf{D}^2 f(x_0)(x_1)](x_2)$.*

Esta definición tiene sentido, pues $\mathbf{D}^2 f(x_0) : \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^m)$ y entonces $\mathbf{D}^2 f(x_0)(x_1) \in L(\mathbb{R}^n, \mathbb{R}^m)$; por lo tanto, se puede aplicar a x_2. La razón por la que hacemos esto es que B_{x_0} evita el uso innecesario del espacio conceptualmente difícil $L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^m))$.

Por definición, una ***aplicación bilineal*** $B : E \times F \to G$, donde E, F, G son espacios vectoriales, es una aplicación que es lineal en cada variable por separado; por ejemplo, en la primera variable esto quiere decir que $B(\alpha e_1 + \beta e_2, f) = \alpha B(e_1, f) + \beta B(e_2, f)$, donde $e_1, e_2 \in E, f \in F$ y $\alpha, \beta \in \mathbb{R}$. Se puede verificar que la aplicación B_{x_0} que hemos definido es una aplicación bilineal de $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$.

Dada una aplicación bilineal $B : E \times F \to \mathbb{R}$, podemos asociarle una matriz con cada elección de bases $e_1, \ldots, e_n$ de E y $f_1, \ldots, f_m$ de F; es decir, podemos hacer

$$a_{ij} = B(e_i, f_j).$$

Si

$$x = \sum_{i=1}^{n} x_i e_i \quad \text{e} \quad y = \sum_{j=1}^{m} y_j f_j,$$

entonces

$$B(x, y) = \sum_{i,j} a_{ij} x_i y_j = (x_1, \ldots, x_n) \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Obsérvese que la matriz asociada a B depende de la elección de las bases de E y F.

Nota. Para la diferencial segunda, haremos un abuso de notación y seguiremos escribiendo $\mathbf{D}^2 f(x_0)$ para la aplicación bilineal B_{x_0} obtenida al diferenciar $\mathbf{D}f$ en x_0 como se describió en los párrafos anteriores.

6.8.2 Teorema *Sea $f: A \subset \mathbb{R}^n \to \mathbb{R}$ dos veces diferenciable en el conjunto abierto A. Entonces la matriz de $\mathbf{D}^2 f(x): \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ en la base canónica está dada por*

$$\begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}$$

donde cada derivada parcial está evaluada en el punto $x = (x_1, \ldots, x_n)$.

Para las derivadas de orden superior, procedemos de forma análoga. Por ejemplo, $\mathbf{D}^3 f$ da una aplicación trilineal $\mathbf{D}^3 f(x): \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ para cada x. No asociamos una matriz con esta aplicación, sino las derivadas de orden tres etiquetadas con cuatro índices: $\partial^3 f_k / \partial x_i \partial x_j \partial x_l$ para cada componente f_k (tales cantidades están muy relacionadas con los llamados ***tensores***).

Antes de pasar al teorema de Taylor, daremos una propiedad importante de la segunda derivada. *La matriz del teorema 6.8.2 es, en condiciones muy débiles, simétrica.*

6.8.3 Simetría de las derivadas parciales cruzadas *Sea $f: A \to \mathbb{R}^m$ dos veces diferenciable en el conjunto abierto A con $\mathbf{D}^2 f$ continua (es decir, con las funciones $\partial^2 f / \partial x_i \partial x_j$ continuas). Entonces $\mathbf{D}^2 f$ es simétrica; es decir,*

$$\mathbf{D}^2 f(x)(x_1, x_2) = \mathbf{D}^2 f(x)(x_2, x_1),$$

o, en términos de las componentes,

$$\frac{\partial^2 f_k}{\partial x_i \partial x_j} = \frac{\partial^2 f_k}{\partial x_j \partial x_i}.$$

Usando esto, se sigue que todas las derivadas de orden superior son simétricas también en condiciones análogas.

La simetría de las segundas derivadas representa una propiedad fundamental que no aparece en el cálculo de una variable. Verifiquemos la simetría en un ejemplo específico. Sea $f(x, y, z) = e^{xy} \operatorname{sen} x + x^2y^4 \cos^2 z$, de modo que $f: \mathbb{R}^3 \to \mathbb{R}$. Entonces

$$\frac{\partial f}{\partial x} = e^{xy} \cos x + ye^{xy} \operatorname{sen} x + 2xy^4 \cos^2 z,$$

$$\frac{\partial f}{\partial y} = xe^{xy} \operatorname{sen} x + 4x^2y^3 \cos^2 z,$$

y

$$\frac{\partial^2 f}{\partial y \partial x} = xe^{xy} \cos x + e^{xy} \operatorname{sen} x + xye^{xy} \operatorname{sen} x + 8xy^3 \cos^2 z = \frac{\partial^2 f}{\partial x \partial y}.$$

El teorema **6.8.3** no es intuitivamente obvio. Sin embargo, se puede obtener cierta intuición a partir de la idea principal de la demostración. Para las funciones $f(x, y)$, consideramos la cantidad

$$S = f(x + h, y + k) - f(x + h, y) - f(x, y + k) + f(x, y),$$

que es la *diferencia de las diferencias* en la figura 6.8-1. El hecho algebraico de que S se puede escribir como la diferencia de las diferencias de dos formas (horizontal y después vertical, o viceversa) es la razón de la igualdad de las derivadas parciales cruzadas.

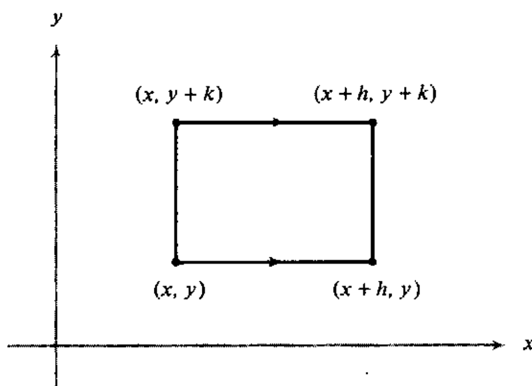


FIGURA 6.8-1 S es una diferencia de diferencias

6.8.4 Definición *Una función es* ***de clase*** C^r *si las primeras r derivadas existen y son continuas (de forma equivalente, esto significa que todas las derivadas parciales*

hasta de orden r existen y son continuas). Una función es **suave** *o* **de clase** C^∞ *si es de clase* C^r *para todos los enteros positivos r.*

El uso iterado de la regla de la cadena se puede usar para demostrar que la composición de funciones C^r también es C^r.

6.8.5 Teorema de Taylor *Sea* $f: A \to \mathbb{R}$ *de clase* C^r *para* $A \subset \mathbb{R}^n$ *un conjunto abierto. Sean x, y* $\in$ *A y supóngase que el segmento de recta que une x con y está contenido en A. Entonces, existe un punto c en ese segmento tal que*

$$f(y) - f(x) = \sum_{k=1}^{r-1} \frac{1}{k!}\mathbf{D}^k f(x)(y - x, \ldots, y - x) + \frac{1}{r!}\mathbf{D}^r f(c)(y - x, \ldots, y - x)$$

donde $\mathbf{D}^k f(x)(y-x, \ldots, y-x)$ *denota* $\mathbf{D}^k f(x)$ *como una aplicación k-lineal aplicada a la k-upla* $(y-x, \ldots, y-x)$. *En coordenadas,*

$$\mathbf{D}^k f(x)(y - x, \ldots, y - x) = \sum_{i_1,\ldots,i_k=1}^{n} \left(\frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}\right)(y_{i_1} - x_{i_1}) \cdots (y_{i_k} - x_{i_k}).$$

Haciendo $y = x + h$, *podemos escribir la fórmula de Taylor como*

$$f(x + h) = f(x) + \mathbf{D}f(x) \cdot h + \cdots + \frac{1}{(r-1)!}\mathbf{D}^{r-1}f(x) \cdot (h, \ldots, h) + R_{r-1}(x, h)$$

donde $R_{r-1}(x, h)$ *es el resto. Además,*

$$\frac{R_{r-1}(x,h)}{\|h\|^{r-1}} \to 0 \quad \text{cuando} \quad h \to 0.$$

En la demostración del teorema se dan otras formas de expresar el término del resto. Este teorema es una generalización del teorema del valor medio (en el caso $r = 1$) y del teorema de Taylor para el cálculo en una variable.

Nota. Las dos últimas afirmaciones del teorema de Taylor sólo requieren que f sea C^{r-1}, como se verá al analizar la demostración.

La idea básica de la demostración del teorema de Taylor es usar el teorema fundamental del cálculo (§**4.8**) para escribir

$$f(x+h)-f(x)=\int_0^1 \frac{d}{dt}f(x+th)\,dt$$

y después integrar por partes varias veces para generar una serie.

El teorema de Taylor nos lleva a construir la ***serie de Taylor*** alrededor de x_0,

$$\sum_{k=0}^{\infty}\frac{1}{k!}\mathbf{D}^k f(x_0)(x-x_0,\ldots,x-x_0).$$

Esta serie no tiene por qué converger a $f(x)$, incluso si f es C^∞. Si converge a f en una vecindad de x_0, decimos que f es ***analítica en sentido real*** en x_0. Así, una función f es analítica en el sentido real si el término del resto $(1/r!)\,\mathbf{D}^r f(c)(x-x_0,\ldots,x-x_0)\to 0$ cuando $r\to\infty$. Por ejemplo, si e^x, sen x y cos x se definen de manera más tradicional y no mediante el enfoque de series de potencias que hemos usado en §**5.4**, entonces se puede utilizar este método para establecer sus desarrollos en serie de potencias.

6.8.6 Ejemplo *Justifíquese y verifíquese después la igualdad de las derivadas parciales cruzadas para $f(x, y) = yx^2(\cos y^2)$.*

Solución Como f es C^∞, el teorema **6.8.3** garantiza la igualdad de las derivadas parciales cruzadas. Para verlo explícitamente, calculamos lo siguiente:

$$\frac{\partial f}{\partial x}=2xy\cos y^2,\quad \frac{\partial^2 f}{\partial y\partial x}=2x\cos y^2-4xy^2\operatorname{sen}y^2;$$

$$\frac{\partial f}{\partial y}=x^2\cos y^2-2y^2x^2\operatorname{sen}y^2,\quad \frac{\partial^2 f}{\partial x\partial y}=2x\cos y^2-4y^2x\operatorname{sen}y^2.\quad \blacklozenge$$

6.8.7 Ejemplo *Si f es C^∞ en $\mathbb{R}$ y para cada intervalo $[a, b]$ existe M constante tal que $\left|f^{(n)}(x)\right|\le M^n$ para todo n y $x\in[a,b]$, muéstrese que f es analítica en cada x_0 y que*

$$f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n.$$

Solución Elegimos b tal que $-b<x_0<b$. El resto es

$$\left|\frac{f^{(n)}(c)}{n!}(x-x_0)^n\right|\le\frac{M^n|x-x_0|^n}{n!},$$

donde c está entre x y x_0 y $\left|f^{(n)}(c)\right| < M^n$ en $[-b, b]$. Esto tiende a 0 cuando $n \to \infty$, ya que, por el criterio del cociente, la serie correspondiente converge. Obsérvese que la convergencia es uniforme en todo intervalo acotado (¿por qué?). ♦

6.8.8 Ejemplo *Dése un ejemplo de una función C^∞ que no sea analítica.*

Solución Sea

$$f(x) = 0 \quad \text{si } x = 0,$$

y

$$f(x) = e^{-1/x^2} \quad \text{si } x \neq 0.$$

El único punto donde la suavidad está en duda es $x = 0$. En ese punto, podemos usar la regla de l'Hôpital para evaluar la derivada como el límite de un cociente de diferencias:

$$f'(0) = \lim_{x\to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x\to 0} \frac{1}{x} e^{-1/x^2} = 0$$

(proporciónense los detalles de la aplicación de la regla de l'Hôpital). Análogamente, si $x \neq 0$, tenemos $f'(x) = (2/x^3)\, e^{-1/x^2}$ y una aplicación análoga de la regla de l'Hôpital indica que $f''(0)$ existe y es 0. Se procede por inducción para mostrar que $f^{(n)}(0)$ existe y es 0 para cada $n > 0$. Así, f es C^∞, pero la serie de Taylor de f con centro en $x_0 = 0$ es idénticamente 0. Esta serie de Taylor no converge al valor de la función en ninguna vecindad no trivial de 0, por lo que f no es analítica en 0. ♦

6.8.9 Ejemplo *Calcúlese la fórmula de Taylor de segundo orden para $f(x, y) = \cos(x + 2y)$ en torno a $(0, 0)$.*

Solución En este caso, $f(0, 0) = 1$,

$$\frac{\partial f}{\partial x}(0,0) = -\operatorname{sen}(0 + 2\cdot 0) = 0, \qquad \frac{\partial f}{\partial y}(0,0) = -2\operatorname{sen}(0 + 2\cdot 0) = 0,$$

$$\frac{\partial^2 f}{\partial x^2}(0,0) = -\cos(0) = -1, \qquad \frac{\partial^2 f}{\partial y^2}(0,0) = -4\cos 0 = -4,$$

y

$$\frac{\partial^2 f}{\partial x\,\partial y}(0,0) = \frac{\partial^2 f}{\partial y\,\partial x}(0,0) = -2\cos 0 = -2.$$

Así,

$$f(h,k) = 1 - \frac{1}{2}(h^2 + 4hk + 4k^2) + R_2((0,0),(h,k)),$$

donde

$$\frac{R_2((0,0),(h,k))}{\|(h,k)\|^2} \to 0 \quad \text{cuando} \quad (h,k) \to (0,0). \quad \blacklozenge$$

Ejercicios de §6.8

1. Verifíquese la igualdad de las derivadas parciales cruzadas para $f(x,y) = (e^{x^2+y^2x})xy^2$.

2. Úsese el ejemplo **6.8.7** para establecer la serie de Taylor y la analiticidad de e^x, sen x y cos x en todo $\mathbb{R}$, suponiendo que conocemos sus derivadas.

3. Sea $f : \mathbb{R} \to \mathbb{R}$ dada por

$$f(x) = x^2 \operatorname{sen}\left(\frac{1}{x}\right) \qquad \text{si } x \in \,]-1,1[,\ x \neq 0$$

y

$$f(x) = 0 \qquad \text{si } x = 0.$$

Analícese la validez del teorema de Taylor para f en torno al punto $x = 0$.

4. Encuéntrese la representación en serie de Taylor en torno a $x = 0$ de $\log(1 - x)$, $-1 < x < 1$, y muéstrese que es igual a $\log(1 - x)$ en $-1 < x < 1$; muéstrese también que converge uniformemente en los subintervalos cerrados de $]-1, 1[$.

5. Calcúlese la fórmula de Taylor de segundo orden para $f(x, y) = e^x \cos y$ en torno a $(0, 0)$.

6. Verifíquese que si se cumplen las condiciones del ejemplo **6.8.7**, entonces podemos derivar la serie de Taylor término a término para obtener $f'(x)$.

§6.9 Máximos y mínimos

Una aplicación importante del teorema de Taylor es la determinación de máximos y mínimos de funciones escalares. Como cabe esperar dado el conocimiento que tenemos de las funciones de una variable, los criterios están relacionados con la segunda derivada. Recordemos en primer lugar el caso de una variable.

Si f : $]a, b[\to \mathbb{R}$ tiene un máximo o mínimo local en x_0 y f es derivable en x_0, entonces $f'(x_0) = 0$. Además, si f es dos veces derivable continuamente y $f''(x_0) < 0$, entonces x_0 es un máximo local y si $f''(x_0) > 0$, entonces x_0 es un mínimo local. Para generalizar estos resultados al caso de funciones escalares de n variables, comenzaremos dando las definiciones pertinentes.

6.9.1 Definición *Sea $f: A \subset \mathbb{R}^n \to \mathbb{R}$ donde A es abierto. Si existe una vecindad de $x_0 \in A$ en la que $f(x_0)$ es un máximo, es decir, si $f(x_0) \geq f(x)$ para todo x en la vecindad, decimos que x_0 es un punto de* ***máximo local*** *y $f(x_0)$ es un valor* ***máximo local*** *de f. Análogamente podemos definir un* ***mínimo local*** *de f. Un punto es* ***extremo*** *si es un máximo o un mínimo local de f. Un punto x_0 es un* ***punto crítico*** *si f es diferenciable en x_0 y si $\mathbf{D}f(x_0) = 0$.*

El primer hecho básico aparece en el siguiente teorema.

6.9.2 Teorema *Si $f: A \subset \mathbb{R}^n \to \mathbb{R}$ es diferenciable, A es abierto y $x_0 \in A$ es un punto extremo de f, entonces $\mathbf{D}f(x_0) = 0$; es decir, x_0 es un punto crítico.*

La demostración es muy parecida a la del caso del cálculo de una variable. El resultado es intuitivamente claro, pues en un punto extremo la gráfica de f debe tener un plano tangente horizontal. Sin embargo, el solo hecho de ser un punto crítico no es suficiente para garantizar que el punto sea extremo. Por ejemplo, considérese $f(x) = x^3$. Para esta función, 0 es un punto crítico, pues $\mathbf{D}f(0) = 0$. Pero $x^3 > 0$ para $x > 0$ y $x^3 < 0$ para $x < 0$, de modo que 0 *no* es extremo. Otro ejemplo está dado por $f(x, y) = y^2 - x^2$. En este caso, $0 = (0, 0)$ es un punto crítico, pues $\partial f/\partial x = -2x$, $\partial f/\partial y = 2y$ y entonces $\mathbf{D}f(0, 0) = 0$. Sin embargo, en cualquier vecindad de 0 podemos encontrar puntos donde f es mayor que 0 y puntos donde f es menor que 0. Un punto crítico que no es un punto extremo local es un ***punto silla***. La figura 6.9-1 muestra el origen de esta terminología.

Para $f: A \subset \mathbb{R} \to \mathbb{R}$, ya hemos mencionado que $f(x)$ es un valor máximo local si $f'(x) = 0$ y $f''(x) < 0$. Esto está claro desde un punto de vista geométrico, si recordamos que $f''(x) < 0$ significa que f es cóncava hacia abajo (o que las pendientes $f'(x)$ son *decrecientes*). Para generalizar esto, presentamos el concepto de hessiana de una función g en x_0.

6.9.3 Definición *Si $g: B \subset \mathbb{R}^n \to \mathbb{R}$ es de clase C^2, la* ***hessiana*** *de g en x_0 se define como la función bilineal $H_{x_0}(g): \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ dada por $H_{x_0}(g)(x, y) = \mathbf{D}^2 g(x_0)(x, y)$. Así, la hessiana es, como matriz, simplemente la matriz de las segundas derivadas parciales.*

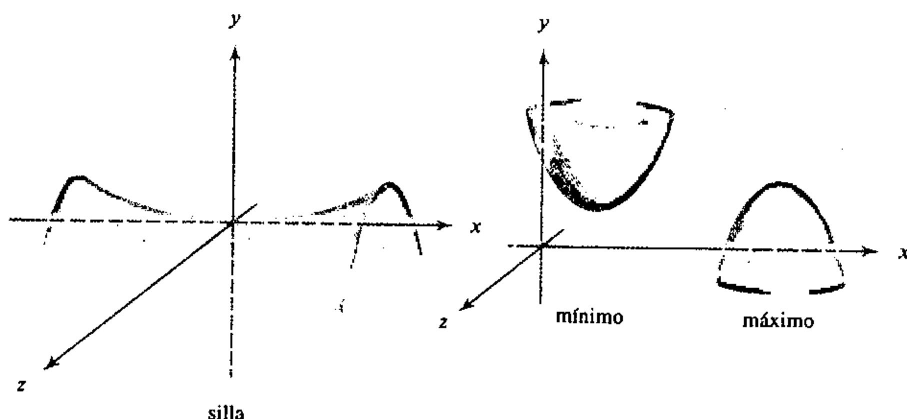


FIGURA 6.9-1 Puntos máximo, mínimo y silla

Una forma bilineal, es decir, una aplicación bilineal $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, es ***definida positiva*** *si $B(x, x) > 0$ para todo $x \neq 0$ en $\mathbb{R}^n$ y es* ***semidefinida positiva*** *si $B(x, x) \geq 0$ para todo $x \in \mathbb{R}^n$. Las formas bilineales* ***definidas negativas*** *y* ***semidefinidas negativas*** *se definen de forma análoga.*

Ahora podemos hacer la siguiente generalización al caso de varias variables.

6.9.4 Teorema

i. *Si $f : A \subset \mathbb{R}^n \to \mathbb{R}$ es una función C^2 definida en un conjunto abierto A y x_0 es un punto crítico de f tal que $H_{x_0}(f)$ es definida negativa, entonces f tiene un máximo local en x_0.*

ii. *Si f tiene un máximo local en x_0, entonces $H_{x_0}(f)$ es semidefinida negativa.*

Para el caso del mínimo se reemplaza "negativa" por "positiva" en el teorema **6.9.4**. Obsérvese que un mínimo de f es un máximo de $-f$.

El polinomio de Taylor de segundo orden define la función cuadrática que mejor aproxima a la gráfica de la función. El criterio de la diferencial segunda dice entonces que podemos determinar el comportamiento en un punto crítico si el término de segundo orden no es degenerado. Es el mismo que el de la superficie cuadrática correspondiente: un paraboloide que se abre hacia arriba o hacia abajo, o una superficie de silla de montar (hiperboloide de una hoja). Los dos primeros corresponden a los casos en

que la hessiana es definida positiva o definida negativa, y el tercero al caso en que no es definida.

Como hemos observado, la matriz de $H_{x_0}(f)$ en la base canónica es

$$\begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1\,\partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_1\,\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n\,\partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_n\,\partial x_n} \end{pmatrix}$$

donde las derivadas parciales se evalúan en x_0.

Si $n = 1$, el teorema **6.9.4i** se reduce al criterio $f''(x_0) < 0$ para una variable; si $f''(x_0) = 0$, se puede tener un máximo, un mínimo o un punto silla (en este caso, el criterio falla). Por ejemplo, $f(x) = -x^4$ tiene un máximo, x^5 un punto silla, y x^4 un mínimo en $x_0 = 0$, aunque $f''(0) = 0$.

Algunos resultados del álgebra lineal serán de utilidad al usar el teorema **6.9.4**. Sea Δ_k el determinante de la matriz

$$\begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1\,\partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_1\,\partial x_k} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_k\,\partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_k\,\partial x_k} \end{pmatrix}.$$

Ésta es la matriz hessiana con las últimas $n - k$ filas y columnas eliminadas. *La matriz simétrica $H_{x_0}(f)$ es definida positiva sii $\Delta_k > 0$ para $k = 1, \ldots, n$.* No demostraremos esto en general aquí, pero en el ejemplo **6.9.5** lo demostraremos para las matrices 2×2. También existe un criterio para el caso definido negativo, dado en el siguiente párrafo. *Así, si $\Delta_k > 0$ para $k = 1, \ldots, n$, entonces f tiene un mínimo (local) en el punto crítico x_0.*

Si una matriz simétrica es semidefinida positiva, entonces $\Delta_k \geq 0$, por lo que **6.9.4ii** implica que si $\Delta_k < 0$ para algún k, f no puede tener un máximo en x_0. Análogamente, f tiene un mínimo (local) en x_0 si $H_{x_0}(f)$ es definida negativa. Al cambiar el signo de $H_{x_0}(f)$ en el párrafo anterior y usar las propiedades de los determinantes, tenemos que $H_{x_0}(f)$ es definida negativa sii $\Delta_k < 0$ para k impar y $\Delta_k > 0$ para k par, y que si $H_{x_0}(f)$ es semidefinida negativa, entonces $\Delta_k \leq 0$ para k impar y $\Delta_k \geq 0$ para k par. Así, f tiene un mínimo en x_0 si $\Delta_k < 0$ para k impar y $\Delta_k > 0$ para k par. Si $\Delta_k > 0$ para algún k impar o $\Delta_k < 0$ para algún k par, entonces f no puede tener un valor mínimo en x_0. De hecho, si $\Delta_k < 0$ para algún k par, f no puede tener ni un máximo ni un mínimo en x_0, y x_0 debe ser un punto silla de f (véase el ejercicio 8 al final de este capítulo para un ejemplo de este caso).

Nota. Este teorema también es útil en economía para la optimización de cantidades como la ganancia. También se usa en mecánica cuando f es el potencial de un sistema, pues un mínimo corresponde a la estabilidad y los máximos y puntos silla corresponden a la inestabilidad (véanse más detalles en Marsden y Tromba, *Cálculo vectorial*, 3ª edición, Addison-Wesley Iberoamericana, 1991, capítulo 4).

6.9.5 Ejemplo *Muéstrese que la matriz*

$$\begin{pmatrix} a & b \\ b & d \end{pmatrix}$$

es definida positiva sii $a > 0$ *y* $ad - b^2 > 0$.

Solución El hecho de que la matriz sea definida positiva significa que

$$(x, y)\begin{pmatrix} a & b \\ b & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} > 0 \quad \text{si } (x, y) \neq (0, 0);$$

es decir, $ax^2 + 2bxy + dy^2 > 0$. En primer lugar, suponemos que esto es cierto para todo $(x, y) \neq (0, 0)$. Hacemos $y = 0$, $x = 1$ para obtener $a > 0$. Si $y = 1$, tenemos $ax^2 + 2bx + d > 0$ para todo x. Esta función es una parábola con un mínimo (pues $a > 0$) en $2ax + 2b = 0$. Es decir, en $x = -b/a$. Por lo tanto,

$$a\left(-\frac{b}{a}\right)^2 + 2b\left(-\frac{b}{a}\right) + d > 0;$$

es decir, $ad - b^2 > 0$. El recíproco se puede demostrar de la misma forma. ♦

6.9.6 Ejemplo *Analícese la naturaleza del punto crítico* (0, 0) *de*

$$f(x, y) = x^2 - xy + y^2.$$

Solución Las derivadas parciales son

$$\frac{\partial f}{\partial x} = 2x - y, \quad \frac{\partial^2 f}{\partial x^2} = 2, \quad \frac{\partial^2 f}{\partial x \partial y} = -1, \quad \frac{\partial f}{\partial y} = -x + 2y, \quad \frac{\partial^2 f}{\partial y^2} = 2,$$

de modo que la hessiana evaluada en $x = 0$, $y = 0$ es

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

En este caso, $\Delta_1 = 2 > 0$ y $\Delta_2 = 4 - 1 = 3 > 0$ de modo que la hessiana es definida positiva. Así, tenemos un mínimo local. ♦

Ejercicios de §6.9

1. Demuéstrese que

$$\begin{pmatrix} a & b \\ b & d \end{pmatrix}$$

es definida negativa sii $a < 0$ y $ad - b^2 > 0$.

2. Analícese la naturaleza del punto crítico $(0, 0)$ de $f(x, y) = x^2 + 2xy + y^2 + 6$.

3. Analícese la naturaleza del punto crítico $(0, 0, 0)$ de $f(x, y, z) = x^2 + y^2 + 2z^2 + xyz$.

4. (Este ejercicio presupone un cierto conocimiento de álgebra lineal.) Sea A una matriz simétrica. Muéstrese que A es definida positiva si y sólo si los autovalores de A (que existen y son reales, pues A es simétrica) son positivos. ¿Es cierto esto si A no es simétrica?

5. Verifíquese que la matriz

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

cumple $\Delta_i \geq 0$ pero que la matriz *no* es semidefinida.

6. Determínese la naturaleza del punto crítico $(0, 0)$ de la función $x^3 + 2xy^2 - y^4 + x^2 + 3xy + y^2 + 10$.

Demostraciones de los teoremas del capítulo 6

6.1.2 Teorema *Sea A un conjunto abierto en $\mathbb{R}^n$ y supóngase que $f : A \to \mathbb{R}^m$ es diferenciable en x_0. Entonces $\mathbf{D}f(x_0)$ queda determinada de forma única por f.*

Demostración Sean L_1 y L_2 dos transformaciones lineales que satisfagan las condiciones de la definición de la derivada. Debemos mostrar que $L_1 = L_2$. Fíjese $e \in \mathbb{R}^n$, $\|e\| = 1$ y sea $x = x_0 + \lambda e$ para $\lambda \in \mathbb{R}$. Entonces, obsérvese que

$$|\lambda| = \|x - x_0\| \quad \text{y} \quad \|L_1 \cdot e - L_2 \cdot e\| = \frac{\|L_1 \cdot \lambda e - L_2 \cdot \lambda e\|}{|\lambda|}.$$

Como A es abierto, $x \in A$ para λ suficientemente pequeño. Por la desigualdad triangular,

$$\begin{aligned} \|L_1 \cdot e - L_2 \cdot e\| &= \frac{\|L_1(x - x_0) - L_2(x - x_0)\|}{\|x - x_0\|} \\ &\leq \frac{\|f(x) - f(x_0) - L_1(x - x_0)\|}{\|x - x_0\|} \\ &\quad + \frac{\|f(x) - f(x_0) - L_2(x - x_0)\|}{\|x - x_0\|}. \end{aligned}$$

Cuando $\lambda \to 0$, estos dos términos tienden a 0, de modo que $L_1 \cdot e = L_2 \cdot e$. Nuestra selección de e fue arbitraria, excepto porque $\|e\| = 1$. Pero para cualquier vector distinto de cero $y \in \mathbb{R}^n$, $y/\|y\| = e$ tiene longitud 1, y, por linealidad, si $L_1(e) = L_2(e)$, entonces $L_1(y) = L_2(y)$. ■

6.2.2 Teorema *Supóngase que $A \subset \mathbb{R}^n$ es un conjunto abierto y que $f: A \to \mathbb{R}^m$ es diferenciable en A. Entonces las derivadas parciales $\partial f_j/\partial x_i$ existen y la matriz de la transformación lineal $\mathbf{D}f(x)$ con respecto a las bases canónicas en $\mathbb{R}^n$ y $\mathbb{R}^m$ está dada por*

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots & \dfrac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \dfrac{\partial f_m}{\partial x_2} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix}$$

*donde cada derivada parcial se evalúa en $x = (x_1, \ldots, x_n)$. Esta matriz es la **matriz jacobiana** de f o **matriz derivada**.*

Demostración Por definición de la matriz de una transformación lineal, el elemento j, i de la matriz de $\mathbf{D}f(x)$ es la j-ésima componente del vector $\mathbf{D}f(x) \cdot e_i = \mathbf{D}f(x)$

aplicada al i-ésimo vector e_i de la base canónica. Llamamos a esta componente a_{ji}. Sea $y = x + he_i$ para $h \in \mathbb{R}$ y obsérvese que

$$\frac{\|f(y) - f(x) - \mathbf{D}f(x)\cdot(y-x)\|}{\|y-x\|} = \frac{\|f(x_1,\ldots,x_i+h,\ldots,x_n) - f(x_1,\ldots,x_n) - h\mathbf{D}f(x)\cdot e_i\|}{|h|}.$$

Como esto tiende a cero cuando $h \to 0$, también lo hace la j-ésima componente del numerador, y entonces, cuando $h \to 0$,

$$\frac{|f_j(x_1,\ldots,x_i+h,\ldots,x_n) - f_j(x_1,\ldots,x_n) - ha_{ji}|}{|h|} \to 0.$$

Por lo tanto,

$$a_{ji} = \lim_{h\to 0} \frac{f_j(x_1,\ldots,x_i+h,\ldots,x_n) - f_j(x_1,\ldots,x_n)}{h} = \frac{\partial f_j}{\partial x_i}. \quad ■$$

6.3.1 Proposición *Supóngase que $A \subset \mathbb{R}^n$ es abierto y que $f : A \to \mathbb{R}^m$ es diferenciable en A. Entonces f es continua. De hecho, para cada $x_0 \in A$ existe una constante $M > 0$ y un $\delta_0 > 0$ tales que $\|x - x_0\| < \delta_0$ implica $\|f(x) - f(x_0)\| \le M\|x - x_0\|$ (esto recibe el nombre de* ***propiedad local de Lipschitz****).*

Para la demostración, recordemos que si $L : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal, entonces existe M_0 constante tal que $\|Lx\| \le M_0\|x\|$ para todo $x \in \mathbb{R}^n$ (véase el ejemplo resuelto **4.4** al final del capítulo 4). Aquí tomaremos $L = \mathbf{D}f(x_0)$.

Demostración Para demostrar la continuidad, basta demostrar la propiedad de Lipschitz ya enunciada, pues dado $\varepsilon > 0$, podemos elegir $\delta = \text{mín}(\delta_0, \varepsilon/M)$. Para demostrar la propiedad de Lipschitz, sea $\varepsilon = 1$ en la definición de la diferencial. Entonces existe δ_0 tal que $\|x - x_0\| < \delta_0$ implica

$$\|f(x) - f(x_0) - \mathbf{D}f(x_0)(x - x_0)\| \le \|x - x_0\|,$$

lo que implica a su vez que

$$\|f(x) - f(x_0)\| \le \|\mathbf{D}f(x_0)(x - x_0)\| + \|x - x_0\|$$

(aquí usamos la desigualdad triangular en la forma $\|y\| - \|z\| \le \|y - z\|$, que se sigue de escribir $y = (y - z) + z$ y aplicar la forma usual de la desigualdad triangular). Sea $M = M_0 + 1$ y usemos el hecho de que $\|\mathbf{D}f(x_0)(x - x_0)\| \le M_0\|x - x_0\|$ para obtener el resultado. ■

6.4.1 Teorema *Sean $A \subset \mathbb{R}^n$ un conjunto abierto y $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$. Supóngase que $f = (f_1, \dots, f_m)$. Si cada derivada parcial $\partial f_j/\partial x_i$ existe y es continua en A, entonces f es diferenciable en A.*

Demostración Si $\mathbf{D}f(x)$ existe, su representación matricial debe ser la matriz jacobiana, por el teorema **6.2.2.** Necesitamos mostrar que con $x \in A$ fijo, para cada $\varepsilon > 0$ existe $\delta > 0$ tal que $\|y - x\| < \delta$, $y \in A$ implica

$$\|f(y) - f(x) - \mathbf{D}f(x)(y - x)\| < \varepsilon\|y - x\|.$$

Para ello, basta demostrarlo para cada componente de f por separado (¿por qué?). Por lo tanto, podemos suponer que $m = 1$. Escribimos

$$\begin{aligned} f(y) - f(x) = {} & f(y_1, \dots, y_n) - f(x_1, y_2, \dots, y_n) + f(x_1, y_2, \dots, y_n) \\ & - f(x_1, x_2, y_3, \dots, y_n) + f(x_1, x_2, y_3, \dots, y_n) \\ & - f(x_1, x_2, x_3, y_4, \dots, y_n) \\ & + \cdots + f(x_1, \dots, x_{n-1}, y_n) - f(x_1, \dots, x_n). \end{aligned}$$

Ahora utilizamos el teorema del valor medio, que implica

$$f(y_1, \dots, y_n) - f(x_1, y_2, \dots, y_n) = \frac{\partial f}{\partial x_1}(u_1, y_2, \dots, y_n)(y_1 - x_1)$$

para algún u_1 entre x_1 e y_1 ($y_2, \dots, y_n$ están fijos). Escribimos expresiones similares para los demás términos y obtenemos

$$\begin{aligned} f(y) - f(x) = {} & \left(\frac{\partial f}{\partial x_1}(u_1, y_2, \dots, y_n)\right)(y_1 - x_1) \\ & + \left(\frac{\partial f}{\partial x_2}(x_1, u_2, y_3, \dots, y_n)\right)(y_2 - x_2) \\ & + \cdots + \left(\frac{\partial f}{\partial x_n}(x_1, x_2, \dots, x_{n-1}, u_n)\right)(y_n - x_n). \end{aligned}$$

Como $\mathbf{D}f(x)(y - x) = \sum_{i=1}^n (\partial f/\partial x_1)(x_1, \dots, x_n)(y_i - x_i)$,

$$\begin{aligned} \|f(y) - f(x) - \mathbf{D}f(x)(y - x)\| \leq {} & \left\{\left|\frac{\partial f}{\partial x_1}(u_1, y_2, \dots, y_n) - \frac{\partial f}{\partial x_1}(x_1, \dots, x_n)\right|\right. \\ & \left. + \cdots + \left|\frac{\partial f}{\partial x_n}(x_1, \dots, x_{n-1}, u_n) - \frac{\partial f}{\partial x_n}(x_1, \dots, x_n)\right|\right\} \|y - x\| \end{aligned}$$

usando la desigualdad triangular y el hecho de que $|y_i - x_i| \leq \|y - x\|$. Como los términos $\partial f/\partial x_i$ son continuos y u_i está entre y_i y x_i, existe $\delta > 0$ tal que el término entre llaves es menor que ε para $\|y - x\| < \delta$. Esta estimación demuestra nuestra afirmación. ■

6.5.1 Regla de la cadena *Sean $A \subset \mathbb{R}^n$ abierto y $f: A \to \mathbb{R}^m$ diferenciable en $x_0 \in A$. Sean $B \subset \mathbb{R}^n$ abierto, $f(A) \subset B$ y $g: B \to \mathbb{R}^p$ diferenciable en $f(x_0)$. Entonces, la composición $g \circ f$ es diferenciable en x_0 y $\mathbf{D}(g \circ f)(x_0) = \mathbf{D}g(f(x_0)) \circ \mathbf{D}f(x_0)$.*

Demostración Para demostrar que $\mathbf{D}(g \circ f)(x_0) \cdot y = \mathbf{D}g(f(x_0)) \cdot (\mathbf{D}f(x_0) \cdot y)$, mostraremos que

$$\lim_{x \to x_0} \frac{\|g \circ f(x) - g \circ f(x_0) - \mathbf{D}g(f(x_0)) \cdot [\mathbf{D}f(x_0)(x - x_0)]\|}{\|x - x_0\|} = 0.$$

Para ello estimamos el numerador como sigue:

$$\begin{aligned}
&\|g \circ f(x) - g \circ f(x_0) - \mathbf{D}g(f(x_0)) \cdot (\mathbf{D}f(x_0)(x - x_0))\| \\
&\quad = \|g(f(x)) - g(f(x_0)) - \mathbf{D}g(f(x_0))[f(x) - f(x_0)] \\
&\quad\quad + \mathbf{D}g(f(x_0))[f(x) - f(x_0) - \mathbf{D}f(x_0)(x - x_0)]\| \\
&\quad \le \|g(f(x)) - g(f(x_0)) - \mathbf{D}g(f(x_0))[f(x) - f(x_0)]\| \\
&\quad\quad + \|\mathbf{D}g(f(x_0))[f(x) - f(x_0) - \mathbf{D}f(x_0)(x - x_0)]\|
\end{aligned}$$

por la desigualdad triangular. Como f es diferenciable, existen δ_0 y $M > 0$ tales que $\|f(x) - f(x_0)\| \le M\|x - x_0\|$ siempre que $\|x - x_0\| < \delta_0$, por el teorema **6.3.1**. Dado $\varepsilon > 0$, por la definición de la diferencial de g, existe $\delta_1 > 0$ tal que $\|y - f(x_0)\| < \delta_1$ implica

$$\|g(y) - g(f(x_0)) - \mathbf{D}g(f(x_0))[y - f(x_0)]\| < \left(\frac{\varepsilon}{2M}\right) \|y - f(x_0)\|.$$

Así, $\|x - x_0\| < \delta_2 = \min\{\delta_0, \delta_1\}$ implica

$$\frac{\|g(f(x)) - g(f(x_0)) - \mathbf{D}g(f(x_0))[f(x) - f(x_0)]\|}{\|x - x_0\|} < \frac{\varepsilon}{2}.$$

Como $\mathbf{D}g(f(x_0))$ es una transformación lineal, sabemos que existe N constante tal que $\|\mathbf{D}g(f(x_0))(y)\| \le N \cdot \|y\|$ para todo $y \in \mathbb{R}^m$, donde podemos suponer que $N \neq 0$. Ahora bien, por la definición de la diferencial, existe $\delta_3 > 0$ tal que $\|x - x_0\| < \delta_3$ implica

$$\frac{\|f(x) - f(x_0) - \mathbf{D}f(x_0)(x - x_0)\|}{\|x - x_0\|} < \frac{\varepsilon}{2N}.$$

Entonces $\|x - x_0\| < \delta_3$ implica

$$\frac{\|\mathbf{D}g(f(x_0))[f(x) - f(x_0) - \mathbf{D}f(x_0)(x - x_0)]\|}{\|x - x_0\|} \le \frac{N\|f(x) - f(x_0) - \mathbf{D}f(x_0)(x - x_0)\|}{\|x - x_0\|} < \frac{\varepsilon}{2}.$$

Sea $\delta = \min\{\delta_2, \delta_3\}$. Así, $\|x - x_0\| < \delta$ implica

$$\frac{\|g \circ f(x) - g \circ f(x_0) - \mathbf{D}g(f(x_0)) \cdot \mathbf{D}f(x_0)(x - x_0)\|}{\|x - x_0\|}$$

$$\leq \frac{\|g(f(x)) - g(f(x_0)) - \mathbf{D}g(f(x_0))[f(x) - f(x_0)]\|}{\|x - x_0\|}$$

$$+ \frac{\|\mathbf{D}g(f(x_0))[f(x) - f(x_0) - \mathbf{D}f(x_0)(x - x_0)]\|}{\|x - x_0\|} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

lo que demuestra la fórmula. ■

> **Nota.** Llegados hasta aquí sería conveniente que el estudiante regresase a la demostración de la regla de la cadena en **§4.7** y comparase ambas demostraciones.

6.6.1 Proposición *Sean $A \subset \mathbb{R}^n$ abierto y $f : A \to \mathbb{R}^m$ y $g : A \to \mathbb{R}$ funciones diferenciables. Entonces gf es diferenciable, y para $x \in A$, $\mathbf{D}(gf)(x) : \mathbb{R}^n \to \mathbb{R}^m$ está dada por $\mathbf{D}(gf)(x) \cdot e = g(x)(\mathbf{D}f(x) \cdot e) + (\mathbf{D}g(x) \cdot e)f(x)$ para todo $e \in \mathbb{R}^n$ (esto tiene sentido, pues $g(x) \in \mathbb{R}$ y $\mathbf{D}g(x) \cdot e \in \mathbb{R}$).*

Demostración Dado $\varepsilon > 0$ y $x_0 \in A$, sea $\delta > 0$ tal que $\|x - x_0\| < \delta$ implique

i. $|g(x)| \leq |g(x_0)| + 1 = M$.

ii. $\|f(x) - f(x_0) - \mathbf{D}f(x_0)(x - x_0)\| \leq \dfrac{\varepsilon}{3M}\|x - x_0\|$.

iii. $\|g(x) - g(x_0) - \mathbf{D}g(x_0)(x - x_0)\| \leq \dfrac{\varepsilon}{3\|f(x_0)\|}\|x - x_0\|$.

iv. $\|g(x) - g(x_0)\| \leq \varepsilon/3M$, donde $\|\mathbf{D}f(x_0)y\| \leq M\|y\|$ (**iii** y **iv** se necesitan solamente si $f(x_0) \neq 0$ y $\mathbf{D}f(x_0) \neq 0$). ¿Por qué es posible esta elección de δ?

Con $\|x - x_0\| < \delta$ y la desigualdad triangular, obtenemos

$$\begin{aligned}
&\|g(x)f(x) - g(x_0)f(x_0) - g(x_0)\mathbf{D}f(x_0)(x - x_0) - [\mathbf{D}g(x_0)(x - x_0)]\, f(x_0)\| \\
&\quad\leq \|g(x)f(x) - g(x)f(x_0) - g(x)\mathbf{D}f(x_0)(x - x_0)\| \\
&\qquad + \|g(x)\mathbf{D}f(x_0)(x - x_0) - g(x_0)\mathbf{D}f(x_0)(x - x_0)\| \\
&\qquad + \|g(x)f(x_0) - g(x_0)f(x_0) - [\mathbf{D}g(x_0)(x - x_0)]\, f(x_0)\| \\
&\quad\leq M \cdot \frac{\varepsilon\|x - x_0\|}{3M} + \frac{\varepsilon}{3M} M\|x - x_0\| + \frac{\varepsilon\|x - x_0\|}{3\|f(x_0)\|} \cdot \|f(x_0)\| \\
&\quad= \varepsilon\|x - x_0\|. \quad \blacksquare
\end{aligned}$$

6.7.1 Teorema del valor medio

i. *Supóngase que $f : A \subset \mathbb{R}^n \to \mathbb{R}$ es diferenciable en un conjunto abierto A. Para cualesquiera $x, y \in A$ tales que el segmento de recta que une x con y esté en A (lo que no debe necesariamente ocurrir para todo x, y), existe un punto c en dicho segmento tal que*

$$f(y) - f(x) = \mathbf{D}f(c) \cdot (y - x).$$

ii. *Supóngase que $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en el conjunto abierto A. Supóngase que el segmento de recta que une x con y está contenido en A y que $f = (f_1, \ldots, f_m)$. Entonces existen puntos $c_1, \ldots, c_m$ en dicho segmento tales que*

$$f_i(y) - f_i(x) = \mathbf{D}f_i(c_i)(y - x), \quad i = 1, \ldots, m.$$

Demostración

i. La función $h : [0, 1] \to \mathbb{R}$ definida por $h(t) = f((1 - t)x + ty$ es derivable en t en $]0, 1[$. Existe $t_0 \in\ ]0, 1[$ tal que $h(1) - h(0) = h'(t_0)(1 - 0)$, por el teorema del valor medio para una variable. Ahora, $h(1) = f(y)$ y $h(0) = f(x)$. Al derivar mediante la regla de la cadena, obtenemos $h'(t_0) = \mathbf{D}f((1 - t_0)x + t_0y)(y - x)$, pues la derivada de $(1 - t)x + ty$ con respecto de t es $y - x$ (justifíquese). Por lo tanto, podemos tomar $c = (1 - t_0)x + t_0y$.

ii. Esto se sigue de aplicar **i** a cada componente de f por separado. $\blacksquare$

6.8.2 Teorema *Sea $f: A \subset \mathbb{R}^n \to \mathbb{R}$ dos veces diferenciable en el conjunto abierto A. Entonces la matriz de $\mathbf{D}^2 f(x) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ en la base canónica está dada por*

$$\begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}$$

donde cada derivada parcial está evaluada en el punto $x = (x_1, \ldots, x_n)$.

Demostración La representación matricial de $\mathbf{D}f : A \to L(\mathbb{R}^n, \mathbb{R})$ es el vector fila $(\partial f/\partial x_1, \ldots, \partial f/\partial x_n)$, de modo que por el teorema **6.2.2**, en una versión adecuada para los vectores fila, $\mathbf{D}^2 f : A \to \mathbb{R}^{n^2}$ está dada por

$$\begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}.$$

Si consideramos $\mathbf{D}^2 f$ como una aplicación bilineal, la representación matricial no cambia, como muestra un análisis de la definición. ■

6.8.3 Simetría de las derivadas parciales cruzadas *Sea $f: A \to \mathbb{R}^m$ dos veces diferenciable en el conjunto abierto A con $\mathbf{D}^2 f$ continua (es decir, con las funciones $\partial^2 f/\partial x_i \partial x_j$ continuas). Entonces $\mathbf{D}^2 f$ es simétrica; es decir,*

$$\mathbf{D}^2 f(x)(x_1, x_2) = \mathbf{D}^2 f(x)(x_2, x_1),$$

o, en términos de las componentes,

$$\frac{\partial^2 f_k}{\partial x_i \partial x_j} = \frac{\partial^2 f_k}{\partial x_j \partial x_i}.$$

Demostración Mostrar que $\mathbf{D}^2 f(x) \cdot (y, z) = \mathbf{D}^2 f(x) \cdot (z, y)$ es equivalente a establecer que

$$\frac{\partial^2 f_k}{\partial x_i \partial x_j} = \frac{\partial^2 f_k}{\partial x_j \partial x_i}.$$

Manteniendo las demás variables fijas, podemos limitarnos al caso bidimensional. Así, podemos suponer que f es de clase C^2 en $A \subset \mathbb{R}^2$ y que es escalar. Como sugiere el texto, consideramos, para $(x, y) \in A$ y h, k pequeños, la cantidad

$$S_{h,k} = [f(x+h, y+k) - f(x+h, y)] - [f(x, y+k) - f(x, y)].$$

Definimos la función g_k como $g_k(u) = f(u, y+k) - f(u, y)$, y, motivados por la figura 6.8-1, observamos que la fórmula para $S_{h,k}$ se puede escribir como

$$S_{h,k} = g_k(x+h) - g_k(x).$$

Así, por el teorema del valor medio, $S_{h,k} = g_k'(c_{h,k}) \cdot h$ para algún $c_{h,k}$ entre x y $x+h$. Por lo tanto,

$$S_{h,k} = \left\{ \frac{\partial f}{\partial x}(c_{h,k}, y+k) - \frac{\partial f}{\partial x}(c_{h,k}, y) \right\} \cdot h = \frac{\partial^2 f}{\partial y\, \partial x}(c_{h,k}, d_{h,k}) \cdot hk$$

para algún $d_{h,k}$ entre y e $y+k$.

Pero $S_{h,k}$ es "simétrica" con respecto de h, k y de x, y. Al intercambiar los dos términos intermedios en $S_{h,k}$, podemos obtener (de la misma forma)

$$S_{h,k} = \frac{\partial^2 f}{\partial x\, \partial y}(\tilde{c}_{h,k}, \tilde{d}_{h,k}) \cdot hk.$$

Al igualar estas dos fórmulas para $S_{h,k}$, cancelar h, k y hacer $h \to 0$, $k \to 0$ (usando la continuidad de $\mathbf{D}^2 f$) obtenemos el resultado. ■

Nota. He aquí un refinamiento de lo anterior: si f es C^1 y $\partial^2 f/\partial x \partial y$ existe y es continua, entonces $\partial^2 f/\partial y \partial x$ existe y ambas son iguales. Esto requiere un poco más de esfuerzo que la demostración anterior, pero la idea es la misma (véase el ejercicio 24 al final de este capítulo).

6.8.5 Teorema de Taylor *Sea $f: A \to \mathbb{R}$ de clase C^r para $A \subset \mathbb{R}^n$ un conjunto abierto. Sean x, $y \in A$ y supóngase que el segmento de recta que une x con y está contenido en A. Entonces existe un punto c en ese segmento tal que*

$$f(y) - f(x) = \sum_{k=1}^{r-1} \frac{1}{k!}\mathbf{D}^k f(x)(y-x, \ldots, y-x) + \frac{1}{r!}\mathbf{D}^r f(c)(y-x, \ldots, y-x)$$

donde $\mathbf{D}^k f(x)(y-x, \ldots, y-x)$ denota $\mathbf{D}^k f(x)$ como una aplicación k-lineal aplicada a la k-upla $(y-x, \ldots, y-x)$. En coordenadas,

$$\mathbf{D}^k f(x)(y-x, \ldots, y-x) = \sum_{i_1, \ldots, i_k = 1}^{n} \left(\frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}} \right) (y_{i_1} - x_{i_1}) \cdots (y_{i_k} - x_{i_k}).$$

Haciendo $y = x + h$, podemos escribir la fórmula de Taylor como

$$f(x+h) = f(x) + \mathbf{D}f(x)\cdot h + \cdots + \frac{1}{(r-1)!}\mathbf{D}^{r-1}f(x)\cdot(h,\ldots,h) + R_{r-1}(x,h)$$

donde $R_{r-1}(x, h)$ es el resto. Además,

$$\frac{R_{r-1}(x,h)}{\|h\|^{r-1}} \to 0 \quad \textit{cuando} \quad h \to 0.$$

Demostración Si recordamos que

$$\frac{d}{dt}f(x+th) = \mathbf{D}f(x+th)\cdot h = \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(x+th)h_i$$

por la regla de la cadena, entonces podemos integrar ambos miembros desde $t = 0$ hasta $t = 1$ para obtener

$$f(x+h) - f(x) = \int_0^1 \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(x+th)h_i\,dt.$$

Si integramos la expresión del miembro derecho por partes y usamos la fórmula general

$$\int_0^1 u\frac{dv}{dt}\,dt = -\int_0^1 v\frac{du}{dt}\,dt + uv\Big|_0^1$$

con $u = (\partial f/\partial x_i)(x+th)h_i$ y $v = t - 1$. Entonces,

$$\sum_{i=1}^{n}\int_0^1 \frac{\partial f}{\partial x_i}(x+th)h_i\,dt = \sum_{i,k=1}^{n}\int_0^1 (1-t)\frac{\partial^2 f}{\partial x_i\,\partial x_k}(x+th)h_ih_k\,dt + \sum_{i=1}^{n}h_i\frac{\partial f}{\partial x_i}(x),$$

pues

$$\frac{du}{dt} = \frac{\partial^2 f}{\partial x_i\,\partial x_k}(x+th)h_ih_k,$$

por la regla de la cadena, y

$$uv\Big|_0^1 = (t-1)\frac{\partial f}{\partial x_i}(x+th)h_i\Big|_{t=0}^{1} = \frac{\partial f}{\partial x_i}(x)h_i.$$

Por lo tanto, hemos demostrado la identidad

$$f(x+h) - f(x) = \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(x)\cdot h_i + R_1(h,x)$$

donde

$$R_1(h,x)=\sum_{i,k=1}^{n}\int_0^1(1-t)\frac{\partial^2 f}{\partial x_i\partial x_k}(x+th)h_ih_k\,dt.$$

Como $|h_i|\leq\|h\|$, obtenemos

$$|R_1(h,x_0)|\leq\|h\|^2\left\{\sum_{i,k=1}^{n}\int_0^1(1-t)\left|\frac{\partial^2 f}{\partial x_i\partial x_k}(x_0+th)\right|dt\right\}$$

Si integramos $R_1(h,x)$ por partes nuevamente, con

$$u=\frac{\partial^2 f}{\partial x_i\partial x_k}(x+th)h_ih_k \quad \text{y} \quad v=-\frac{(t-1)^2}{2},$$

obtenemos

$$R_1(h,x)=\sum_{i,j,k}\int_0^1\frac{(t-1)^2}{2}\frac{\partial^3 f}{\partial x_i\partial x_j\partial x_k}(x+th)h_ih_jh_k\,dt+\sum_{i,j}\frac{1}{2}\frac{\partial^2 f}{\partial x_i\partial x_j}(x)h_ih_j.$$

Así, hemos demostrado que

$$f(x+h)=f(x)+\sum_{i=1}^{n}h_i\frac{\partial f}{\partial x_i}(x)+\sum_{i,j=1}^{n}\frac{h_ih_j}{2}\frac{\partial^2 f}{\partial x_i\partial x_j}(x)+R_2(h,x),$$

donde

$$R_2(h,x)=\sum_{i,j,k}\int_0^1\frac{(t-1)^2}{2}\frac{\partial^3 f}{\partial x_i\partial x_j\partial x_k}(x+th)h_ih_jh_k\,dt.$$

El integrando de la última fórmula es una función continua y, por lo tanto, está acotada en una pequeña vecindad de x (recuérdese que tiene que estar cerca del valor de la función en x). Así, para $M\geq 0$ constante y $\|h\|$ pequeña, obtenemos

$$|R_2(h,x)|\leq\|h\|^3M.$$

En particular, $R_2(h,x_0)/\|h\|^2\leq\|h\|M\to 0$ cuando $h\to 0$. La fórmula para el resto establecida en el teorema (la forma de Lagrange para el resto) se obtiene al aplicar el *segundo teorema del valor medio* para integrales. Recuérdese que éste establece que

$$\int_a^b f(x)g(x)\,dx=f(c)\int_a^b g(x)\,dx,$$

siempre que f y g sean continuas y $g \geq 0$ en $[a, b]$; en este caso, c es cierto número entre a y b (véase el ejercicio 45, capítulo 4). Así,

$$R_1(h, x_0) = \sum_{i,k=1}^{n} \int_0^1 (1 - t) \frac{\partial^2 f}{\partial x_i \partial x_k}(x + th) h_i h_k \, dt$$

$$= \int_0^1 (1 - t) \mathbf{D}^2(f(x + th))(h, h) \, dt = \frac{1}{2} \mathbf{D}^2 f(c) \cdot (h, h)$$

donde c está en la recta que une x con $y = x + h$. Análogamente,

$$R_2(h, x_0) = \sum_{i,j,k=1}^{n} \int_0^1 \frac{(t-1)^2}{2} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(x + th) h_i h_j h_k \, dt$$

$$= \int_0^1 \frac{(t-1)^2}{2} \mathbf{D}^3 f(x + th) \cdot (h, h, h) \, dt = \frac{1}{3!} \mathbf{D}^3 f(c) \cdot (h, h, h)$$

donde c está en la recta que une x con $y = x + h$. Se puede proceder por inducción, usando el mismo método para obtener el resultado general. ■

Observación 1 Con un poco más de esfuerzo, se podría demostrar un teorema más fuerte: si f es de clase C^r, entonces

$$f(x + h) = f(x) + \sum_{k=1}^{r} \frac{1}{k!} \mathbf{D}^k f(x) \cdot (h, \ldots, h) + R_r(x, h)$$

donde

$$\frac{R_r(x, h)}{\|h\|^r} \to 0$$

cuando $h \to 0$ en $\mathbb{R}^n$. Dejaremos esto para el lector interesado.

Observación 2 Otra demostración utiliza la fórmula de Taylor del cálculo de una variable como sigue. Sea $g(t) = f(x + t(y - x))$ para $x \in [0, 1]$. Si aplicamos la fórmula de Taylor en $\mathbb{R}$, existe $\tilde{t} \in [0, 1]$ tal que

$$g(1) - g(0) = \sum_{k=1}^{r-1} \frac{1}{k!} g^{(k)}(0) + \frac{1}{r!} g^{(r)}(\tilde{t}).$$

Obsérvese que $g(1) = f(y)$ y $g(0) = f(x)$. Sea $p(t) = x + t(y - x)$. Entonces $g = f \circ p$ y $\mathbf{D}p(t)(1) = y - x$ para todo x, y en consecuencia, por el ejercicio 6b al final de este capítulo,

$$g^{(r)}(t) = \mathbf{D}^k g(t)(1, \ldots, 1) = \mathbf{D}^k f(p(t))(\mathbf{D}p(t)(1), \ldots, \mathbf{D}p(t)(1))$$
$$= \mathbf{D}^k f(x + t(y - x))(y - x, \ldots, y - x).$$

Sustituyendo, obtenemos

$$f(y)-f(x)=\sum_{k=1}^{r-1}\frac{1}{k!}\mathbf{D}^k f(x)(y-x,\ldots,y-x)+\frac{1}{r!}\mathbf{D}^r f(x+\tilde{t}(y-x))(y-x,\ldots,y-x),$$

lo que completa la demostración, con $c = x + \tilde{t}\,(y-x)$.

6.9.2 Teorema *Si $f: A \subset \mathbb{R}^n \to \mathbb{R}$ es diferenciable, A es abierto y $x_0 \in A$ es un punto extremo de f, entonces $\mathbf{D}f(x_0) = 0$; es decir, x_0 es un punto crítico.*

Demostración Si $\mathbf{D}f(x_0) \neq 0$, podemos encontrar que $x \in \mathbb{R}^n$ tal que $\mathbf{D}f(x_0)x = c \neq 0$; digamos que $c > 0$. Ahora encontramos un $\delta > 0$ tal que si $\|h\| < \delta$, entonces

$$\|f(x_0+h)-f(x_0)-\mathbf{D}f(x_0)h\| \leq \frac{c}{2\|x\|}\|h\|.$$

Sea $\lambda > 0$ tal que $\lambda\|x\| < \delta$. Si elegimos $h = 2x$, obtenemos $\|f(x_0+\lambda x)-f(x_0)-\mathbf{D}f(x_0)\lambda x\| \leq c\lambda\|x\|/(2\|x\|) = c\lambda/2$. Como $\mathbf{D}f(x_0)\lambda x = \lambda c$, obtenemos $f(x_0 + \lambda x) - f(x_0) > 0$. Análogamente, $\|f(x_0-\lambda x)-f(x_0)+\mathbf{D}f(x_0)\lambda x\| \leq c\lambda/2$ implica $f(x_0-\lambda x)-f(x_0) < 0$. Como $f(x_0+\lambda x) > f(x_0)$ y $f(x_0-\lambda x) < f(x_0)$, vemos que $f(x_0)$ no es un valor extremo local. Es decir, podemos determinar puntos y arbitrariamente cercanos a x_0 tales que $f(y) > f(x_0)$ y, análogamente, existen puntos y arbitrariamente cercanos a x_0 tales que $f(y) < f(x_0)$. ■

6.9.4 Teorema

i. *Si $f: A \subset \mathbb{R}^n \to \mathbb{R}$ es una función C^2 definida en un conjunto abierto A y x_0 es un punto crítico de f tal que $H_{x_0}(f)$ es definida negativa, entonces f tiene un máximo local en x_0.*

ii. *Si f tiene un máximo local en x_0, entonces $H_{x_0}(f)$ es semidefinida negativa.*

Demostración

i. $H_{x_0}(f)(x, x) < 0$ para todo $x \neq 0$ en $\mathbb{R}^n$ implica $\mathbf{D}^2f(x_0)(x, x) < 0$ para todo $x \neq 0$ en $\mathbb{R}^n$. Por el ejemplo resuelto **4.5** al final del capítulo 4, una aplicación bilineal es continua, y entonces $\mathbf{D}^2f(x_0)(x, x)$ es una función continua de x. Además, $S = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$ es compacto, por lo que existe un punto $\tilde{x} \in S$ tal que $0 > \mathbf{D}^2f(x_0)(\tilde{x}, \tilde{x}) \geq \mathbf{D}^2f(x_0)(x, x)$ para todo $x \in S$. Sea $\varepsilon = -\mathbf{D}^2f(x_0)(\tilde{x}, \tilde{x})$. Entonces $\mathbf{D}^2f(x_0)(x, x) = \|x\|^2\mathbf{D}^2f(x_0)(x/\|x\|, x/\|x\|) \leq -\varepsilon\|x\|^2$ para todo $x \neq 0$ en $\mathbb{R}^n$. Como $\mathbf{D}^2f$ es continua, existe $\delta > 0$ tal que $\|y - x_0\| < \delta$ implica $\|\mathbf{D}^2f(y) - \mathbf{D}^2f(x_0)\| < \varepsilon/2$, y pode-

mos también elegir δ de modo que $D(x_0, \delta) \subset A$. Si $y \in D(x_0, \delta)$, el teorema de Taylor nos da

$$f(y) - f(x_0) = \mathbf{D}f(x_0)(y - x_0) + \frac{1}{2}\mathbf{D}^2 f(c)(y - x_0, y - x_0),$$

donde $c \in D(x_0, \delta)$. Así,

$$\|\mathbf{D}^2 f(c) - \mathbf{D}^2 f(x_0)\| < \frac{\varepsilon}{2}$$

implica

$$\begin{aligned} \mathbf{D}^2 f(c)(y - x_0, y - x_0) & \\ &\leq \mathbf{D}^2 f(x_0)(y - x_0, y - x_0) \\ &\quad + \|\mathbf{D}^2 f(c)(y - x_0, y - x_0) - \mathbf{D}^2 f(x_0)(y - x_0, y - x_0)\| \\ &\leq -\varepsilon\|y - x_0\|^2 + \frac{\varepsilon}{2}\|y - x_0\|^2 = -\frac{\varepsilon}{2}\|y - x_0\|^2. \end{aligned}$$

Como $\mathbf{D}f(x_0) = 0$, el teorema de Taylor implica

$$f(y) - f(x_0) = \frac{1}{2}\mathbf{D}^2 f(c)(y - x_0, y - x_0) \leq \frac{1}{2}\left(-\frac{\varepsilon}{2}\|y - x_0\|^2\right) < 0.$$

Por lo tanto, $f(y) < f(x_0)$ para todo $y \in D(x_0, \delta)$, $y \neq x_0$ por lo que f tiene un máximo local en x_0.

ii. Para demostrar esta parte del teorema, razonaremos por reducción al absurdo. Sea f una función con un máximo local en x_0 y supóngase que $\mathbf{D}^2 f(x_0)(x, x) > 0$ para algún $x \in \mathbb{R}^n$. Ahora considérese $g(t) = -f(x_0 + tx)$. Como f está definida en una vecindad de x_0, g está definida en una vecindad de 0. Tenemos que $\mathbf{D}^2 g(0)(1, 1) = -\mathbf{D}^2 f(x_0)(x, x) < 0$. Usamos la demostración de **i** para afirmar que existe δ tal que si $|t| < \delta$, $t \neq 0$ entonces $g(t) < g(0)$. Así, $|t| < \delta$ implica $f(x_0 + tx) > f(x_0)$, por lo que f no tiene un máximo local en x_0. Esta contradicción implica que $\mathbf{D}^2 f(x_0)(x, x) \leq 0$ para todo $x \in \mathbb{R}^n$. Por lo tanto, $H_{x_0}(f)(x, x) \leq 0$ para todo $x \in \mathbb{R}^n$. ■

Ejemplos resueltos del capítulo 6

Ejemplo 6.1 *Sea $f: B \subset \mathbb{R}^n \to \mathbb{R}$, donde $B = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$, una función continua que además es diferenciable en* int(B). *Supóngase que $f(x) = 0$ en $\partial(B)$. Muéstrese que existe un punto $x_0 \in$* int(B) *tal que $\mathbf{D}f(x_0) = 0$.*

Solución Ésta es la versión multidimensional del teorema de Rolle. Si f es idénticamente cero, el teorema es trivial. Por lo tanto, supongamos que $f(x) \neq 0$ para algún $x \in \text{int}(B)$. Entonces f alcanza un máximo o un mínimo en un punto interior, pues B es compacto. Así, existe un punto extremo $x_0 \in \text{int}(B)$ y el teorema **6.9.2** implica que $\mathbf{D}f(x_0) = 0$. ♦

Ejemplo 6.2 *Muéstrese que para una aplicación bilineal* $f: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$, *tenemos que* $\mathbf{D}f(x_0, y_0)(x, y) = f(x_0, y) + f(x, y_0)$.

Solución Sabemos que f es diferenciable, pues vemos en su representación matricial que las derivadas parciales son lineales y, por tanto, continuas. Como $f(x, y_0)$ es una función lineal de x, la derivada en la dirección $(x, 0)$ es $\mathbf{D}f(x_0, y_0)(x, 0) = f(x, y_0)$, como en el ejemplo **6.2.4**. Análogamente, $\mathbf{D}f(x_0, y_0)(0, y) = f(x_0, y)$. Como $\mathbf{D}f(x_0, y_0)$ es lineal y $(x, y) = (x, 0) + (0, y)$, tenemos que $\mathbf{D}f(x_0, y_0)(x, y) = f(x_0, y) + f(x, y_0)$. ♦

Ejemplo 6.3 *Encuéntrese la jacobiana de* $f(x, y) = (\text{sen}(x \,\text{sen}\, y), (x + y)^2)$, $f: \mathbb{R}^2 \to \mathbb{R}^2$.

Solución Tenemos que

$$\begin{aligned}
\frac{\partial f_1}{\partial x} &= \frac{\partial}{\partial x}(\text{sen}\,(x \,\text{sen}\, y)) \\
&= \cos(x \,\text{sen}\, y)\frac{\partial}{\partial x}(x \,\text{sen}\, y) = \text{sen}\, y \cos(x \,\text{sen}\, y); \\
\frac{\partial f_2}{\partial x} &= 2(x + y)\frac{\partial}{\partial x}(x + y) = 2(x + y); \\
\frac{\partial f_1}{\partial y} &= \cos(x \,\text{sen}\, y)\frac{\partial}{\partial y}(x \,\text{sen}\, y) = \cos(x \,\text{sen}\, y) x \cos y; \\
\frac{\partial f_2}{\partial y} &= 2(x + y).
\end{aligned}$$

Así, por el teorema **6.2.2**, la matriz jacobiana (donde $x = x_1$ e $y = x_2$) es

$$\begin{pmatrix} \text{sen}\, y \cos(x \,\text{sen}\, y) & x \cos y \cos(x \,\text{sen}\, y) \\ 2(x + y) & 2(x + y) \end{pmatrix}. \quad ♦$$

Nota. Las matrices jacobianas no son simétricas en general; de hecho, ni siquiera tienen que ser cuadradas. La simetría es una propiedad única de la *segunda* derivada de una función $f: \mathbb{R}^n \to \mathbb{R}$.

Ejemplo 6.4 *Determínense los puntos críticos de $f(x, y) = x^3 - 3x^2 + y^2$ y si f tiene un máximo (local), un mínimo (local) o un punto silla en cada uno de estos puntos críticos.*

Solución Los puntos críticos son precisamente aquellos (x, y) para los que

$$\frac{\partial f}{\partial x} = 3x^2 - 6x = 0 \quad \text{y} \quad \frac{\partial f}{\partial y} = 2y = 0.$$

Al despejar x vemos que o $x = 0$ o $x = 2$. Por lo tanto, los puntos críticos de f son (0, 0) y (2, 0). La matriz hessiana en (x, y) es

$$\begin{pmatrix} \dfrac{\partial^2 f}{\partial x \partial x} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y \partial y} \end{pmatrix} = \begin{pmatrix} 6x - 6 & 0 \\ 0 & 2 \end{pmatrix}.$$

En (0, 0), la hessiana es

$$\begin{pmatrix} -6 & 0 \\ 0 & 2 \end{pmatrix}$$

y $\Delta_1 = -6$, $\Delta_2 = -12$. Por lo tanto, f tiene un punto silla en (0, 0). En (2, 0), la hessiana es

$$\begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix}$$

y $\Delta_1 = 6$, $\Delta_2 = 12$, por lo que f tiene un mínimo local. Esto lo confirma con la gráfica por computador que aparece en la figura 6.ER-1.

Ejemplo 6.5 *Sean A un conjunto abierto convexo en $\mathbb{R}^n$ y $f: \mathbb{R}^n \to \mathbb{R}^m$ una función diferenciable, con derivada continua. Supóngase que $\|\mathbf{D}f(x)y\| \leq M\|y\|$ para todo $x \in A$ e $y \in \mathbb{R}^n$. Demuéstrese la* ***desigualdad del valor medio:***

$$\|f(x_1) - f(x_2)\| \leq M\|x_1 - x_2\|.$$

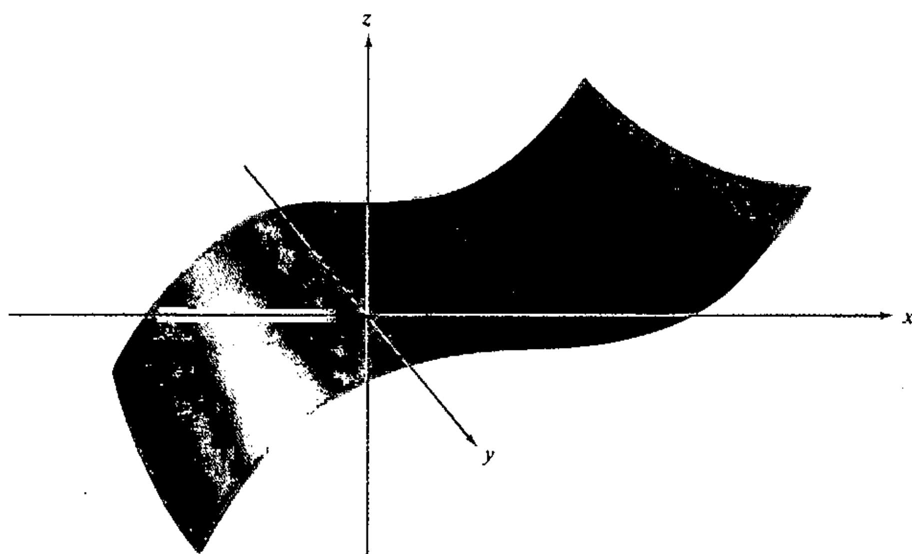


FIGURA 6.ER-1 La gráfica de $f(x, y) = x^3 - 3x^2 + y^2$

Solución Para $n = 1$, $m = 1$, esto se sigue directamente del teorema del valor medio. Para el caso general, obsérvese que, por la regla de la cadena, $(d/dt)\,(f(tx_1 + (1-t)x_2)) = \mathbf{D}f(tx_1 + (1-t)x_2) \cdot (x_1 - x_2)$. Integramos ambos miembros con respecto a t, desde $t = 0$ hasta $t = 1$ para obtener

$$f(x_1) - f(x_2) = \int_0^1 [\mathbf{D}f(tx_1 + (1-t)x_2) \cdot (x_1 - x_2)]\, dt.$$

Esta integral se define como la integral de las funciones componentes. Al adquirir valores absolutos y usar la hipótesis sobre $\mathbf{D}f$ obtenemos el resultado pedido. Hemos usado el hecho de que el valor absoluto de una integral es menor o igual que la integral del valor absoluto (véanse las observaciones después de **4.8.5** y obsérvese que el caso de las funciones vectoriales es análogo). ♦

Ejercicios del capítulo 6

1. Si $f: A \subset \mathbb{R}^n \to \mathbb{R}^m$ y $g: B \subset \mathbb{R}^n \to \mathbb{R}^m$ son funciones diferenciables en los conjuntos (abiertos) A y B y α, β son constantes, demuéstrese que $\alpha f + \beta g : A \cap B \subset \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable y que $\mathbf{D}(\alpha f + \beta g)(x) = \alpha \mathbf{D}f(x) + \beta \mathbf{D}g(x)$.

2. Sea $f: A \subset \mathbb{R} \to \mathbb{R}^m$ y supóngase que df_i/dx existe para $i = 1, \ldots, m$. Muéstrese que $\mathbf{D}f$ existe.

3. Sea $f: [0, \infty[\to \mathbb{R}$ una función continua tal que f es diferenciable en $]0, \infty[$. Supóngase que $f(0) = 0$ y que $f(x) \to 0$ cuando $x \to +\infty$. Muéstrese que existe $c \in]0, \infty[$ tal que $f'(c) = 0$.

4. Si $f: A \subset \mathbb{R}^n \to \mathbb{R}^m$ es una función constante, muéstrese que $\mathbf{D}f(x) = 0$ para todo $x \in A$.

5. Calcúlense las jacobianas de las siguientes funciones:

a. $f(x, y) = \operatorname{sen}(x^2 + y^3)$.

b. $f(x, y, z) = (z \operatorname{sen} x, z \operatorname{sen} y)$.

c. $f(x, y) = xy$.

d. $f(x, y, z) = x^2 + y^2$.

e. $f(x, y) = (\operatorname{sen}(xy), \cos(xy), x^2y^2)$.

f. $f(x, y, z) = x^{y+z}$.

g. $f(x, y, z) = xyz$.

h. $f(x, y, z) = (z^{xy}, x^2, \tan(xyz))$.

6. a. Si $f: A \subset \mathbb{R}^n \to \mathbb{R}^m$ y $g: B \subset \mathbb{R}^m \to \mathbb{R}^p$ son dos veces diferenciables y $f(A) \subset B$, entonces para $x_0 \in A$, $x, y \in \mathbb{R}^n$ muéstrese que

$$\begin{aligned}\mathbf{D}^2(g \circ f)(x_0)(x, y) = {}& \mathbf{D}^2(g(f(x_0)))(\mathbf{D}f(x_0) \cdot x, \mathbf{D}f(x_0) \cdot y) \\ &+ \mathbf{D}g(f(x_0)) \cdot \mathbf{D}^2 f(x_0)(x, y).\end{aligned}$$

b. Si $p: \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal más una constante y $f: A \subset \mathbb{R}^m \to \mathbb{R}^s$ es k veces diferenciable, demuéstrese que

$$\mathbf{D}^k(f \circ p)(x_0)(x_1, \ldots, x_k) = \mathbf{D}^k f(p(x_0))(\mathbf{D}p(x_0)(x_1), \ldots, \mathbf{D}p(x_0)(x_k)).$$

7. Determínense los puntos críticos de las siguientes funciones y si éstos son máximos locales, mínimos locales o puntos silla:

a. $f(x, y) = x^3 + 6x^2 + 3y^2 - 12xy + 9x$.

b. $f(x, y) = \operatorname{sen} x + y^2 - 2y + 1$.

c. $f(x, y, z) = \cos 2x \cdot \operatorname{sen} y + z^2$.

d. $f(x, y, z) = (x + y + z)^2$.

8. Muéstrese que si $f: A \subset \mathbb{R}^2 \to \mathbb{R}$ tiene un punto crítico $x_0 \in A$ y

$$\Delta = \frac{\partial^2 f}{\partial x_1 \partial x_1} \cdot \frac{\partial^2 f}{\partial x_2 \partial x_2} - \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2$$

se evalúa en x_0, entonces

a. $\Delta > 0$ y $\partial^2 f/\partial x_1 \partial x_1 > 0$ implican que f tiene un mínimo local en x_0.

b. $\Delta > 0$ y $\partial^2 f/\partial x_1 \partial x_1 < 0$ implican que f tiene un máximo local en x_0.

c. $\Delta < 0$ implica que x_0 es un punto silla de f.

9. Considérense las siguientes dos propiedades para un subconjunto X de $\mathbb{R}^n$:

1 Existe un punto $x_0 \in X$ tal que cualquier otro punto x de X se puede unir con x_0 mediante una línea recta contenida en X.

2 Existe un punto $x_0 \in X$ tal que cualquier otro punto x en X se puede unir con x_0 mediante una curva diferenciable en X.

a. Dense ejemplos de cada tipo de conjunto que no sean convexos.

b. Muéstrese que si X es un conjunto abierto en $\mathbb{R}^n$ que satisface cualquiera de estas condiciones y $f: X \to \mathbb{R}$ es una función diferenciable con diferencial nula, entonces f es constante.

c. Muéstrese que si X es un subconjunto abierto de $\mathbb{R}^n$, entonces las siguientes propiedades son equivalentes:

i. La condición **2** anterior

ii. La conexión por arcos de X

iii. La conexión de X

10. Demuéstrese el análogo del teorema **6.9.4** para mínimos.

11. Demuéstrese el análogo de la proposición **5.3.3** para $f: A \subset \mathbb{R}^n \to \mathbb{R}^m$.

12. Una función $f: \mathbb{R}^n \to \mathbb{R}$ es ***homógenea de grado*** m si $f(tx) = t^m f(x)$ para todo $x \in \mathbb{R}^n$ y $t \in \mathbb{R}$. Si f es diferenciable, muéstrese que para $x \in \mathbb{R}^n$,

$$\mathbf{D}f(x)x = mf(x), \quad \text{es decir,} \quad \sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i} = mf(x).$$

Muéstrese que las aplicaciones multilineales en k variables dan lugar a funciones homogéneas de grado k. Dense otros ejemplos.

13. Úsese la regla de la cadena para encontrar las derivadas de las siguientes funciones, donde $f(x, y, z) = x^2 + yz$, $g(x, y) = y^3 + xy$ y $h(x) = \text{sen } x$:

* **a.** $F(x, y, z) = f(h(x), g(x, y), z)$.

b. $G(x, y, z) = h(f(x, y, z)g(x, y))$.

c. $H(x, y, z) = g(f(x, y, h(x)), g(z, y))$.

Determínense también fórmulas *generales* para las derivadas de F, G, H.

14. **a.** Extienda el ejemplo resuelto **6.2** a las aplicaciones multilineales.

b. Aplíquese el resultado en **a** al caso de la aplicación determinante $\det : \mathbb{R}^{n^2} = \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$, para mostrar que $A \in \mathbb{R}^{n^2}$ es un punto crítico de det sii A tiene rango $\leq n - 2$.

15. Sea $f : \mathbb{R} \to \mathbb{R}$ diferenciable. Supóngase que *no* existe $x \in \mathbb{R}$ tal que f y f' se anulen *simultáneamente* en x. Muéstrese que $S = \{x \mid 0 \leq x \leq 1, f(x) = 0\}$ es finito.

16. Si $f : \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable y $\mathbf{D}f$ es constante, muéstrese que f es un término lineal más una constante y que la parte lineal de f es el valor constante de $\mathbf{D}f$.

17. Si $f : A \subset \mathbb{R}^n \to \mathbb{R}$ es de clase C^r y $\mathbf{D}f(x_0) = 0, \mathbf{D}^2 f(x_0) = 0, \ldots, \mathbf{D}^{r-1}f(x_0) = 0$ pero $\mathbf{D}^r f(x_0)(x, \ldots, x) < 0$ para todo $x \in \mathbb{R}^n$, $x \neq 0$, demuéstrese entonces que f tiene un máximo local en x_0.

18. Demuéstrese que la ecuación $x^3 + bx + c = 0$ con $b > 0$ tiene exactamente una solución $x \in \mathbb{R}$.

19. En cada uno de los siguientes problemas, determínese la fórmula de Taylor de segundo orden para la función dada en torno al punto dado (x_0, y_0):

* **a.** $f(x, y) = (x + y)^2$, $x_0 = 0, y_0 = 0$.

b. $f(x, y) = e^{x+y}$, $x_0 = 0$, $y_0 = 0$.

c. $f(x, y) = 1/(x^2 + y^2 + 1)$, $x_0 = 0$, $y_0 = 0$.

d. $f(x, y) = e^{-x^2-y^2} \cos(xy)$, $x_0 = 0$, $y_0 = 0$.

e. $f(x, y) = \text{sen } (xy) + \cos(xy)$, $x_0 = 0$, $y_0 = 0$.

f. $f(x, y) = e^{(x-1)^2} \cos y$, $x_0 = 1$, $y_0 = 0$.

20. Sea $L : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Defínase $\|L\| = \inf\{M \mid \|Lx\| \leq M\|x\|$ para todo $x \in \mathbb{R}^n\}$. Muéstrese que $\|\cdot\|$ es una norma en el espacio de las transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$.

21. **a.** Sea $f: A \subset \mathbb{R}^n \to \mathbb{R}$ y supóngase que para algún entero k, $\Delta_k < 0$, donde Δ_k se evalúa en x_0. Muéstrese que f no puede tener un mínimo (local) en x_0.

b. Si $\Delta_k < 0$ para algún k par, demuéstrese que f tiene un punto silla en x_0.

22. Dése un ejemplo de una función continua $f:]0, 1[\to \mathbb{R}$ cuya gráfica no sea cerrada. ¿Puede ocurrir esto para $f: A \subset \mathbb{R} \to \mathbb{R}$, donde A es cerrado?

23. Escríbanse los primeros cuatro términos del desarrollo de Taylor de $\log(\cos x)$ en torno a $x = 0$.

24. Sea $f(x, y)$ una función escalar sobre $\mathbb{R}^2$. Muéstrese que si f es de clase C^1 y $\partial^2 f/\partial x \partial y$ existe y es continua, entonces $\partial^2 f/\partial y \partial x$ existe y $\partial^2 f/\partial x \partial y = \partial^2 f/\partial y \partial x$ (esto es más débil que decir que f sea de clase C^2). Generalícese.

25. Sea $f: \mathbb{R}^n \to \mathbb{R}$ y supóngase que las derivadas parciales $\partial f/\partial x_i = 1, \ldots, n$, existen y que $\partial f/\partial x_i$, $i = 1, \ldots, n-1$ son continuas. Demuéstrese que f es diferenciable.

26. **a.** Si $f: \mathbb{R} \to \mathbb{R}$ y f' existe en una vecindad de $x = a$ y $\lim_{x\to a^+} f'(x) = l$, demuéstrese que $f'(a) = l$.

b. Sea $g(x) = 1$ si $x < 0$ y $g(x) = 0$ si $x \geq 0$. ¿Puede ser g la derivada de alguna función?

27. Sea $f: A \to \mathbb{R}$ continua, donde $A \subset \mathbb{R}^n$ es abierto. Supóngase que todas las derivadas direccionales existen y que en cada $x_0 \in A$ determinan una transformación lineal. ¿Debe f ser diferenciable?

28. Supóngase que definimos un nuevo tipo de conexión como sigue: un conjunto $M \subset \mathbb{R}^n$ es "diferenciablemente conexo" si cualesquiera dos puntos $x, y \in M$ pueden unirse mediante un arco diferenciable $\psi: [0, 1] \to \mathbb{R}^n$. ¿Cómo son los conjuntos diferenciablemente conexos? Dése un ejemplo de conjunto conexo por arcos que no sea diferenciablemente conexo. ¿Qué ocurre con los conjuntos abiertos?

29. Sean $f_n(x) = xe^{-nx}$, $x \in [0, \infty[$, $n = 0, 1, 2, \ldots$.

a. Muéstrese que $f(x) = \sum_{n=0}^{\infty} f_n(x)$ existe. Calcúlese f explícitamente.

b. ¿Es f continua?

c. Encuéntrese un conjunto adecuado donde la convergencia sea uniforme.

d. ¿Podemos derivar término a término?

30. Supóngase que $f: \mathbb{R} \to \mathbb{R}$ está acotada y que tiene derivada continua. ¿Qué es correcto y qué incorrecto en la siguiente serie de argumentos?

Queremos demostrar que el conjunto T de todos los puntos en los que f tiene un máximo (absoluto) es cerrado. Como f es diferenciable, es continua. Por lo tanto, alcanza su máximo; es decir, T no es vacío. Denotamos por S el conjunto de puntos en los que $f'(x) = 0$. Entonces $T \subset S$. Por otro lado, si $x \in S$, entonces $f'(x) = 0$; por lo tanto, f alcanza un máximo o un mínimo. Si alcanza un máximo, debemos tener $f(x) \geq 0$. Por lo tanto, $T = S \cap \{x \mid f(x) \geq 0\}$. Como $\{x \mid f(x) \geq 0\}$ es cerrado, al igual que S, T es cerrado.

¿Es T realmente cerrado, o no lo es?

31. Sea $A \subset \mathbb{R}^n$ compacto, y construyamos el espacio normado $C(A, \mathbb{R})$ como en el capítulo 5. Defínase, para $x_0 \in A$, $\delta_{x_0} : C(A, \mathbb{R}) \to \mathbb{R}$; $f \mapsto f(x_0)$. Demuéstrese que δ_{x_0} es diferenciable.

32. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x, y) = \frac{xy(x^2 - y^2)}{x^2 + y^2}$$

si $(x, y) \neq (0, 0)$ y $f(0, 0) = 0$. Muéstrese que $\partial^2 f/\partial x \partial y$ y $\partial^2 f/\partial y \partial x$ existen en $(0, 0)$ pero que no son iguales.

33. Úsese el teorema de Taylor para demostrar el teorema del binomio,

$$(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

34. Considérese la sucesión de números reales

$$\frac{1}{2}, \frac{1}{2 + \frac{1}{2}}, \frac{1}{2 + \frac{1}{2+\frac{1}{2}}}, \ldots$$

(una fracción continua). Muéstrese que es convergente y determínese su límite.

35. Sea $f :]a, b[\to \mathbb{R}$ dos veces diferenciable. Supóngase que f se anula en tres puntos distintos. Demuéstrese que existe $c \in]a, b[$ tal que $f''(c) = 0$.

36. Supóngase que $f : [0, 1] \to \mathbb{R}$ es continua, f es diferenciable en $]0, 1[$ y que $f(0) = 0$. Supóngase que $|f'(x)| \leq |f(x)|$, $0 < x < 1$. Demuéstrese que $f(x) = 0$ para todo $x \in [0, 1]$.

37. Una función C^2, $f : \mathbb{R}^2 \to \mathbb{R}$, es ***armónica*** si

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0.$$

Supóngase que (x_0, y_0) es un máximo local estricto y que f es armónica. Demuéstrese que *todas* las derivadas segundas de f se anulan en (x_0, y_0).

38. Determínese la ecuación del plano tangente a las siguientes superficies en los puntos indicados:

a. $z = x^2 + y^2$, $(0, 0)$.

b. $z = x^2 - y^2 + x$, $(1, 0)$.

c. $z = (x + y)^2$, $(3, 2)$.

39. Analícese el comportamiento de las siguientes funciones en los puntos indicados:

a. $z = x^2 - y^2 + 3xy$, $(0, 0)$.

b. $z = Ax^2 - By^2 + Cxy$, $(0, 0)$.

40. Encuéntrese la ecuación del plano tangente a la superficie S dada por la gráfica de

a. $f(x, y) = \sqrt{x^2 + y^2} + x^2 + y^2$ en $(1, 0, 2)$.

b. $f(x, y) = \sqrt{x^2 + 2xy - y^2 + 1}$ en $(1, 1, \sqrt{3})$.

41. Dése un ejemplo de función diferenciable $f: \mathbb{R} \to \mathbb{R}$ tal que $f(0) = 0$, $f'(0) = 0$ y que no cumpla una desigualdad de la forma $|f(x)| \leq Cx^2$ en ninguna vecindad del origen.

Capítulo 7
Teoremas de la función inversa e implícita y temas relacionados

En álgebra lineal aprendemos que un sistema de ecuaciones lineales

$$\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & y_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & \cdots & + & a_{nn}x_n & = & y_n \end{array}$$

tiene una única solución $x_1, \ldots, x_n$ si la matriz $A = (a_{ij})$ no es singular; es decir, si $\det(A) \neq 0$, donde $\det(A)$ denota el determinante de A. ¿Qué ocurre con las ecuaciones no lineales? Para un sistema de la forma

$$\begin{aligned} f_1(x_1, \ldots, x_n) &= y_1 \\ &\vdots \\ f_n(x_1, \ldots, x_n) &= y_n \end{aligned}$$

¿cuándo podemos despejar $x_1, \ldots, x_n$? El objeto que generaliza el determinante en el caso de los sistemas lineales es el ***determinante jacobiano***, definido por $Jf(x) = \det(\mathbf{D}f(x))$, donde $x = (x_1, \ldots, x_n)$ y $f = (f_1, \ldots, f_n)$. Escrito en coordenadas,

$$Jf(x) = \begin{vmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial x_1} & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{vmatrix},$$

donde las derivadas parciales se evalúan en $x = (x_1, \ldots, x_n)$. A veces se escribe Jf como

$$\frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)}.$$

Si $Jf(x) \neq 0$ cabe esperar que se pueda despejar x en $f(x) = y$.

El teorema que justifica tales resultados es el tema principal de **§7.1.** Consideraremos el caso en que queremos despejar y en $f(x, y) = 0$ (el teorema de la función implícita) en **§7.2.** En las secciones posteriores de este capítulo aplicaremos a ecuaciones diferenciales ordinarias algunos teoremas de existencia análogos y obtendremos un resultado relativo a la forma de una función cerca de un máximo, mínimo o punto silla, llamado "lema de Morse". La última sección estudia los problemas de extremos condicionados.

§7.1 Teorema de la función inversa

Como $Jf(x)$ es el determinante de $\mathbf{D}f(x)$, $Jf(x) \neq 0$ implica que $\mathbf{D}f(x) : \mathbb{R}^n \to \mathbb{R}^n$ es un isomorfismo lineal (es decir, que su matriz es invertible). Como la mejor aproximación lineal de f es invertible, nos gustaría concluir que la propia función f es invertible. Sin embargo, aparecen algunas restricciones, que apreciaremos si primero consideramos el caso $f : \mathbb{R} \to \mathbb{R}$. Es cierto que si f es C^1 y si $f'(x_0) \neq 0$, entonces f es invertible (inyectiva) en una vecindad de x_0. Geométricamente, esto está claro, pues $f'(x_0) \neq 0$ indica que f tiene una pendiente distinta de cero *en* x_0, y, en consecuencia, *cerca de* x_0 (véase la figura 7.1-1).

Si $f'(x_0) = 0$, entonces f puede o no ser invertible cerca de x_0; en la figura 7.1-1, f *no* es invertible cerca de x_1, pero *es* invertible cerca de x_3, aunque f' se anule en ambos puntos (cerca de x_1 se comporta como $c_1 + (x - x_1)^2$ y cerca de x_3 se comporta como $c_3 + (x - x_3)^3$). Así, si $f'(x_0) = 0$, no se puede obtener una conclusión (se necesitaría un análisis más a fondo). Además, $f'(x_0) \neq 0$ no garantiza que podamos resolver la ecuación $f(x) = y$ *para todo* y. Por ejemplo, no existe x_4 tal que $f(x_4) = y_4$ para y_4 como en la figura 7.1-1. De la misma figura podemos ver que las soluciones no son únicas por lo general, ya que $f(x_0) = f(x_2)$. Existirá una única solución sólo si restringimos nuestra atención a una vecindad adecuadamente pequeña de x_0.

Por lo tanto, lo único que podemos esperar es que f sea invertible *cerca* de $f(x_0)$. Es decir, que para y cercana a $f(x_0)$, podamos encontrar un único x cercano a x_0 tal que $f(x) = y$. La pregunta "¿a qué cercanía?" es muy sutil y requiere un análisis detallado de la demostración. Así, nuestro interés principal será el de la ***invertibilidad local***, es decir, la invertibilidad de $f(x)$ para x cerca de x_0 e y cerca de $f(x_0)$.

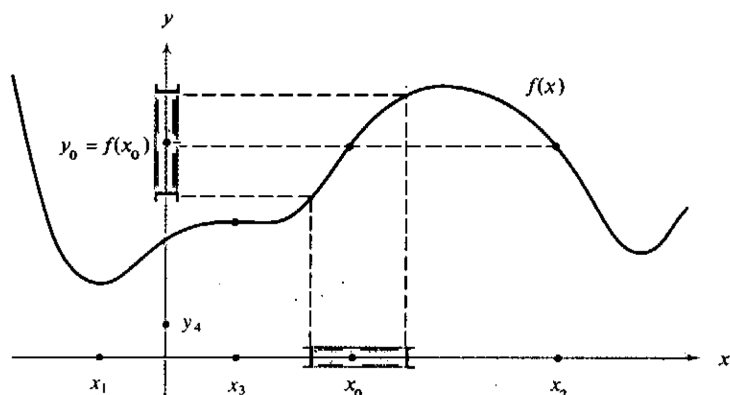


FIGURA 7.1-1 Esta función tiene una inversa cerca de x_0

Para calcular la derivada de la función inversa $f^{-1}(y)$, usamos la regla de la cadena: Como $f^{-1}(f(x)) = x$, obtenemos $(df^{-1}/dy) \cdot f'(x) = 1$ y, por lo tanto,

$$\left.\frac{df^{-1}}{dy}\right|_{y=f(x)} = \frac{1}{df/dx}.$$

Demostrar que f^{-1} es realmente diferenciable requiere un poco más de cuidado. El teorema 7.1.1 incluye la situación de una variable como un caso particular.

7.1.1 Teorema de la función inversa *Sean $A \subset \mathbb{R}^n$ un conjunto abierto y $f : A \subset \mathbb{R}^n \to \mathbb{R}^n$ de clase C^1 (es decir, supóngase que $\mathbf{D}f$ existe y es continua). Sea $x_0 \in A$ y supóngase que $Jf(x_0) \neq 0$. Entonces existen una vecindad U de x_0 en A y una vecindad abierta W de $f(x_0)$, tales que $f(U) = W$ y la restricción de f a U tiene una inversa C^1 $f^{-1}: W \to U$. Además, para $y \in W$ y $x = f^{-1}(y)$, tenemos*

$$\mathbf{D}f^{-1}(y) = [\mathbf{D}f(x)]^{-1},$$

la inversa de $\mathbf{D}f(x)$, lo que significa su inversa como transformación lineal (correspondiente a la matriz inversa). Si f es de clase C^p, $p \geq 1$, también lo es f^{-1}.

Nota.

1. Si $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ es de clase C^1 y $\mathbf{D}f(x_0)$ es inyectiva, entonces f también es localmente inyectiva cerca de x_0. Análogamente, si $\mathbf{D}f(x_0)$ es suprayectiva, entonces también lo es f en una vecindad de $f(x_0)$. Estos resultados se siguen del teorema de la función inversa, por los métodos de **§7.2**; véase el ejercicio 11 al final de este capítulo.

2. Decir que f tiene una inversa f^{-1} significa exactamente que podemos resolver de forma única $f(x) = y$ para $x \in U$, dado cualquier $y \in W$. El hecho de que exista una solución se debe a que f aplica U *sobre* W (suprayectividad) y que la solución sea única se debe a que f es *inyectiva* en U (inyectividad).

Para y cercano a y_0, necesitamos demostrar la *existencia* y *unicidad* de un x tal que $f(x) = y$. La herramienta técnica básica que usaremos es el principio de la aplicación contractiva. En **§5.7** vimos cómo se podía usar este resultado para demostrar la existencia de soluciones de algunas ecuaciones diferenciales e integrales sencillas, lo que ampliaremos en **§7.5**.

7.1.2 Ejemplo *Considérense las ecuaciones* $(x^4 + y^4)/x = u(x, y)$ *y* $\operatorname{sen} x + \cos y = v(x, y)$. *¿En torno a qué puntos* (x, y) *podemos despejar* x, y *en términos de* u, v?

Solución En este caso, las funciones son $u(x, y) = f_1(x, y) = (x^4 + y^4)/x$ y $v(x, y) = f_2(x, y) = \operatorname{sen} x + \cos y$. Queremos determinar los puntos cerca de los cuales podemos despejar x, y como funciones de u y v y calcular $\partial x/\partial v$, etcétera. Para usar el teorema de la función inversa, primero debemos calcular $\partial(f_1, f_2)/\partial(x, y)$. Definiendo $f = (f_1, f_2)$, tomamos su dominio como $A = \{(x, y) \in \mathbb{R}^2 \mid x \neq 0\}$. Entonces

$$\frac{\partial(f_1,f_2)}{\partial(x,y)} = \begin{vmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{vmatrix} = \begin{vmatrix} \dfrac{3x^4 - y^4}{x^2} & \dfrac{4y^3}{x} \\ \cos x & -\operatorname{sen} y \end{vmatrix}$$

$$= \frac{\operatorname{sen} y}{x^2}(y^4 - 3x^4) - \frac{4y^3}{x}\cos x.$$

En los puntos en los que esta expresión está definida y no se anula, podemos despejar x, y en términos de u y v. En otras palabras, podemos despejar x, y cerca de aquellos x, y tales que $x \neq 0$ y $(\operatorname{sen} y)(y^4 - 3x^4) \neq 4xy^3 \cos x$. Por lo general, tales condiciones no se pueden resolver explícitamente. Por ejemplo, si $x_0 = \pi/2$, $y_0 = \pi/2$, podemos despejar x, y cerca de x_0, y_0, pues en este caso, $\partial(f_1, f_2)/\partial(x, y) \neq 0$.

Las derivadas $\partial x/\partial u$, etcétera, se obtienen de acuerdo con el teorema de la función inversa, invirtiendo la matriz jacobiana. En el caso 2×2, esto equivale a lo siguiente:

$$\frac{\partial x}{\partial u} = \frac{1}{Jf(x,y)}\frac{\partial v}{\partial y}, \qquad \frac{\partial x}{\partial v} = \frac{-1}{Jf(x,y)}\frac{\partial u}{\partial y},$$

$$\frac{\partial y}{\partial u} = \frac{-1}{Jf(x,y)}\frac{\partial v}{\partial x}, \qquad \frac{\partial y}{\partial v} = \frac{1}{Jf(x,y)}\frac{\partial u}{\partial x}$$

como podemos comprobar multiplicando

$$\frac{1}{Jf}\begin{pmatrix} \dfrac{\partial v}{\partial y} & \dfrac{-\partial u}{\partial y} \\ \dfrac{-\partial v}{\partial x} & \dfrac{\partial u}{\partial x} \end{pmatrix}\begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{pmatrix} = \frac{1}{Jf}\begin{pmatrix} \dfrac{\partial v}{\partial y}\dfrac{\partial u}{\partial x} - \dfrac{\partial u}{\partial y}\dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y}\dfrac{\partial u}{\partial y} - \dfrac{\partial u}{\partial y}\dfrac{\partial v}{\partial y} \\ -\dfrac{\partial v}{\partial x}\dfrac{\partial u}{\partial x} + \dfrac{\partial u}{\partial x}\dfrac{\partial v}{\partial x} & -\dfrac{\partial v}{\partial x}\dfrac{\partial u}{\partial y} + \dfrac{\partial u}{\partial x}\dfrac{\partial v}{\partial y} \end{pmatrix}$$

$$= \frac{1}{Jf}\begin{pmatrix} Jf & 0 \\ 0 & Jf \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

(véase también el ejemplo resuelto **7.2** al final del capítulo).

En este caso

$$\frac{\partial x}{\partial u} = \frac{-(x^2 \operatorname{sen} y)}{(\operatorname{sen} y)(y^4 - 3x^4) - 4y^3 x \cos x}.$$

Obsérvese que la respuesta se expresa en términos de x, y y no de u, v. La evaluación en un punto (x, y) da el valor $\partial x/\partial u$ en el punto correspondiente $(u(x, y), v(x, y))$. En nuestro ejemplo, $x_0 = y_0 = \pi/2$, de donde obtenemos $u = \pi^3/4$ y $v = 1$. La última fórmula se convierte en

$$\left.\frac{\partial x}{\partial u}\right|_{u=\pi^3/4, v=1} = \left.\frac{-x^2 \operatorname{sen} y}{(\operatorname{sen} y)(y^4 - 3x^4) - 4y^3 x \cos x}\right|_{x=y=\pi/2} = \frac{2}{\pi^2}. \quad \blacklozenge$$

El teorema de la función inversa es útil, pues nos dice que existen soluciones de ecuaciones y nos explica cómo derivar las soluciones, aunque sea imposible resolver dichas ecuaciones explícitamente.

7.1.3 Ejemplo *Sean* $u(x, y) = e^x \cos y$ *y* $v(x, y) = e^x \operatorname{sen} y$. *Muéstrese que la transformación* $(x, y) \mapsto (u(x, y), v(x, y))$ *es localmente invertible cerca de todo punto, pero que no es invertible.*

Solución En este caso,

$$\frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} e^x \cos y & -e^x \operatorname{sen} y \\ e^x \operatorname{sen} y & e^x \cos y \end{vmatrix}$$
$$= e^{2x}(\cos^2 y + \operatorname{sen}^2 y) = e^{2x} \neq 0.$$

Por lo tanto, el teorema de la función inversa implica que la aplicación es localmente invertible. Sin embargo, no es globalmente inyectiva, pues

$$u(x, y + 2\pi) = u(x, y), \quad v(x, y + 2\pi) = v(x, y). \quad \blacklozenge$$

Obsérvese que para $f : \mathbb{R} \to \mathbb{R}$, si f es diferenciable y si $f'(x) \neq 0$ para todo x, entonces $f'(x) > 0$ o < 0, pues f' satisface el teorema de los valores intermedios (véase el ejercicio 37, capítulo 4); por lo tanto, f debe ser (globalmente) inyectiva, pues f es siempre creciente o decreciente. El ejemplo **7.1.3** muestra que éste no es necesariamente el caso en $\mathbb{R}^2$.

Ejercicios de §7.1

1. Sea $u(x, y) = x^2 - y^2$, $v(x, y) = 2xy$. Muéstrese que la aplicación $(x, y) \mapsto (u, v)$ es localmente invertible en todo punto $(x, y) \neq (0, 0)$.

2. Calcúlese $\partial x/\partial u$, $\partial x/\partial v$, $\partial y/\partial u$, $\partial y/\partial v$ para las funciones del ejercicio 1.

3. Sea $f(x) = x + 2x^2 \operatorname{sen}(1/x)$, $x \neq 0$, $f(0) = 0$. Muéstrese que $f'(0) \neq 0$, pero que f no es localmente invertible cerca de 0. ¿Por qué no contradice esto el teorema de la función inversa?

4. Sea $L : \mathbb{R}^n \to \mathbb{R}^n$ un isomorfismo lineal y $f(x) = L(x) + g(x)$, donde $\|g(x)\| \leq M\|x\|^2$ y f es C^1. Muéstrese que f es localmente invertible cerca de 0.

5. Analícese si del sistema

$$u(x, y, z) = x + xyz,$$
$$v(x, y, z) = y + xy,$$
$$w(x, y, z) = z + 2x + 3z^2$$

se pueden despejar x, y, z en términos de u, v, w cerca de $(0, 0, 0)$.

§7.2 Teorema de la función implícita

En el estudio del teorema de la función implícita, nuevamente nos interesa la existencia y diferenciabilidad de funciones definidas implícitamente. Sin duda, el estudiante ya ha trabajado con funciones definidas de forma implícita; sin embargo, tal vez no se le ha explicado adecuadamente la justificación de las manipulaciones. Las posibles preguntas que nos gustaría plantear serán más evidentes después de analizar algunos ejemplos. Considérense aquellos x e y relacionados mediante la ecuación $F(x, y) = 0$. Quisiéramos decir que esto define una función $y = f(x)$ (se dice que $y = f(x)$ ***está definida de forma implícita***), y quisiéramos calcular dy/dx. Como en la sección anterior, dada tal F, por lo general no se puede despejar y explícitamente, por lo que es importante saber que tal función existe y saber cómo hacer cálculos con ella sin tener que despejarla.

Como motivación para el teorema, considérese la función $F(x, y) = x^2 + y^2 - 1$. Nos interesan los x e y relacionados por $F(x, y) = 0$, que es precisamente la circunferencia unidad. Una función $f(x)$ es una "solución" cuando $F(x, f(x)) = 0$ para todo x en el dominio de f. Está claro que f está dada por los dos valores $f(x) = \pm\sqrt{1 + x^2}$, y cualquiera de ellos es una solución, así que f *no tiene por qué ser única*. Dado (x_0, y_0) tal que $F(x_0, y_0) = 0$, deseamos saber si podemos encontrar $f(x)$ tal que $F(x, f(x)) = 0$ y f sea diferenciable y *única cerca de* (x_0, y_0). Si $x_0 \neq \pm 1$, esto es cierto si f se toma como la raíz cuadrada adecuada. El valor y_0 dado determina la raíz cuadrada que debe seleccionarse. Véase la figura 7.2-1. Los puntos $x_0 = \pm 1$ son excepcionales por varias razones. En primer lugar, f no es diferenciable en ellos; en segundo lugar, cerca de $x_0 = \pm 1$, f podría ser cualquiera de las raíces cuadradas, por lo que no quedaría determinada de forma única. Estos puntos excepcionales son aquellos para los que $\partial F/\partial y = 0$. Así, en general, queremos una condición como $\partial F/\partial y \neq 0$ para garantizar que, al menos localmente, *podemos* encontrar una única f diferenciable tal que $F(x, f(x)) = 0$.

En el caso general, consideremos una función $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ y estudiemos la relación $F(x, y) = 0$ o, desarrollada,

$$F_1(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0$$
$$\vdots$$
$$F_m(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0.$$

El objetivo es despejar las m incógnitas $y_1, \ldots, y_m$ de las m ecuaciones en términos de $x_1, \ldots, x_n$.

7.2.1 Teorema de la función implícita *Sean $A \subset \mathbb{R}^n \times \mathbb{R}^m$ un conjunto abierto y $F : A \to \mathbb{R}^m$ una función de clase C^p (es decir, supóngase que F tiene p diferenciales continuas, donde p es un entero, $p \geq 1$). Supóngase que $(x_0, y_0) \in A$ y $F(x_0, y_0) = 0$.*

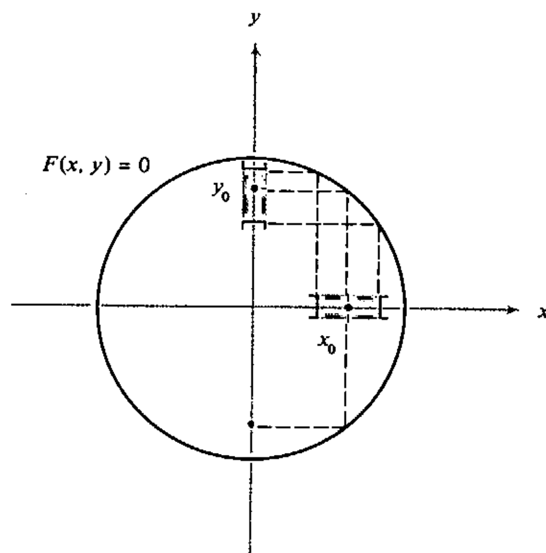


FIGURA 7.2-1 El círculo define de forma implícita dos funciones

Fórmese el determinante

$$\Delta = \begin{vmatrix} \dfrac{\partial F_1}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F_m}{\partial y_1} & \cdots & \dfrac{\partial F_m}{\partial y_m} \end{vmatrix}$$

evaluado en (x_0, y_0), *donde* $F = (F_1, \ldots, F_m)$. *Supóngase que* $\Delta \neq 0$. *Entonces existen una vecindad abierta* $U \subset \mathbb{R}^n$ *de* x_0, *una vecindad* V *de* y_0 *en* $\mathbb{R}^m$ *y una única función* $f\colon U \to V$ *tal que*

$$F(x, f(x)) = 0$$

para todo $x \in U$. *Además,* f *es de clase* C^p.

Demostraremos este teorema utilizando el teorema de la función inversa aplicado a la transformación $(x, y) \mapsto (x, F(x, y))$. La razón intuitiva de la validez del teorema y la necesidad de la restricción $\Delta \neq 0$ debe quedar clara del ejemplo de la figura 7.2-1.

De la ecuación $F(x, f(x)) = 0$, se puede determinar $\mathbf{D}f$ mediante la regla de la cadena. En primer lugar, sea $m = 1$. Entonces, la regla de la cadena implica

$$0 = \frac{\partial}{\partial x_i} F(x, f(x)) = \frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial y}\frac{\partial f}{\partial x_i},$$

por lo que obtenemos la importante ecuación (obsérvese el signo menos):

$$\frac{\partial f}{\partial x_i} = -\frac{\partial F/\partial x_i}{\partial F/\partial y}.$$

Una advertencia especial para el lector es que, debido al signo menos en la expresión

$$\frac{\partial F/\partial x_i}{\partial F/\partial y}$$

es *incorrecto* "cancelar" las ∂F para obtener $\partial y/\partial x$. Así, aunque en ocasiones es útil este recurso nemotécnico, tiene sus limitaciones.

Podemos formular la solución general análoga al teorema **7.2.1.**

7.2.2 Corolario *En el teorema de la función implícita, $\partial f_j/\partial x_i$ está dado por*

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix} = -\begin{pmatrix} \dfrac{\partial F_1}{\partial x_1} & \cdots & \dfrac{\partial F_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F_m}{\partial x_1} & \cdots & \dfrac{\partial F_m}{\partial x_m} \end{pmatrix}^{-1} \begin{pmatrix} \dfrac{\partial F_1}{\partial x_1} & \cdots & \dfrac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F_m}{\partial x_1} & \cdots & \dfrac{\partial F_m}{\partial x_n} \end{pmatrix},$$

donde $^{-1}$ denota la matriz inversa.

La demostración es similar al caso anterior $m = 1$ y se deja como ejercicio.

7.2.3 Ejemplo *Considérese el siguiente sistema de ecuaciones:*

$$xu + yv^2 = 0$$
$$xv^3 + y^2u^6 = 0.$$

¿Se pueden despejar unas únicas u, v en términos de x, y cerca de $x = 1$, $y = -1$, $u = 1$, $v = -1$? ¿Y cerca de $x = 0$, $y = 1$, $u = 0$, $v = 0$? Calcúlese $\partial u/\partial x$ en $x = 1$, $y = -1$ y en $x = 0$, $y = 1$, si existe.

Solución Escribimos las ecuaciones como $F(x, y, u, v) = 0$, donde F representa los miembros izquierdos de las ecuaciones dadas. Queremos ver si podemos despejar $u(x, y)$, $v(x, y)$. Así, formamos

$$\Delta = \begin{vmatrix} \dfrac{\partial F_1}{\partial u} & \dfrac{\partial F_1}{\partial v} \\ \dfrac{\partial F_2}{\partial u} & \dfrac{\partial F_2}{\partial v} \end{vmatrix} = \begin{vmatrix} x & 2yv \\ 6y^2u^5 & 3xv^2 \end{vmatrix}.$$

En $x = 1$, $y = -1$, $u = 1$, $v = -1$, obtenemos

$$\Delta = \begin{vmatrix} 1 & 2 \\ 6 & 3 \end{vmatrix} = -9 \neq 0,$$

por lo que existe una única solución u, v cerca de este punto. Al derivar las ecuaciones dadas con respecto a x obtenemos el par de ecuaciones

$$u + x\frac{\partial u}{\partial x} + 2yv\frac{\partial v}{\partial x} = 0$$

y

$$v^3 + 3xv^2\frac{\partial v}{\partial x} + 6y^2u^5\frac{\partial u}{\partial x} = 0,$$

evaluando en el punto dado, se obtiene el resultado

$$1 + \frac{\partial u}{\partial x} + 2\frac{\partial v}{\partial x} = 0$$

$$-1 + 6\frac{\partial u}{\partial x} + 3\frac{\partial v}{\partial x} = 0.$$

Si eliminamos $\partial v/\partial x$ para despejar $\partial u/\partial x$ se obtiene $\partial u/\partial x = 5/9$. En $x = 0$, $y = 1$, $u = 0$, $v = 0$, tenemos $\Delta = 0$. Así, el teorema de la función implícita establece que no podemos *esperar* tener una solución única u, v en términos de x, y en este caso. Para determinar realmente la solubilidad tendríamos que hacer un análisis directo, no proporcionado por el teorema de la función implícita. ♦

Ejercicios de §7.2

1. Verifíquense de forma directa los puntos cerca de los cuales se puede despejar y en términos de x en la ecuación $F(x, y) = y^2 + y + 3x + 1 = 0$.

2. Verifíquese que la respuesta que se dé al ejercicio 1 coincide con la respuesta que se espera según el teorema de la función implícita. Calcúlese dy/dx.

3. En el sistema

$$\begin{aligned} 3x + 2y + z^2 + u + v^2 &= 0 \\ 4x + 3y + z + u^2 + v + w + 2 &= 0 \\ x + z + w + u^2 + 2 &= 0, \end{aligned}$$

analícese la solubilidad de u, v, w en términos de x, y, z cerca de $x = y = z = 0$, $u = v = 0$, $w = -2$.

4. ¿Tiene la aplicación

$$(x, y) \mapsto \left(\frac{x^2 - y^2}{x^2 + y^2}, \frac{xy}{x^2 + y^2} \right)$$

una inversa local cerca de (0, 1)?

5. Analícese la solubilidad de

$$\begin{aligned} y + x + uv &= 0 \\ uxy + v &= 0 \end{aligned}$$

para u, v en términos de x, y cerca de $x = y = u = v = 0$, y verifíquese directamente.

§7.3 Teorema de rectificación del dominio

Ahora daremos una consecuencia del teorema de la función implícita, que es una importante herramienta técnica en el estudio de las superficies. Este resultado establece, *grosso modo*, que si $f : A \subset \mathbb{R}^n \to \mathbb{R}$ tiene una derivada distinta de cero en un punto x_0, entonces, en una vecindad de x_0, f se puede "rectificar"; de hecho, f se puede deformar hasta convertirse en la aplicación dada por la proyección sobre el eje de coordenadas x_n, mediante su composición con un "cambio de coordenadas", que es (por definición) una función suave con inversa suave. Véase la figura 7.3-1, donde el cambio de coordenadas, que se denota h, rectifica las superficies con f constante en planos. El resultado preciso se establece en el siguiente teorema.

7.3.1 Teorema de rectificación del dominio—*Sean $A \in \mathbb{R}^n$ un conjunto abierto y $f : A \to \mathbb{R}$ una función de clase C^p, $p \geq 1$. Sea $x_0 \in A$ y supóngase que $f(x_0) = 0$ y que* $\mathbf{D}f(x_0) \neq 0$. *Entonces existen un conjunto abierto U, un conjunto abierto V que contiene a x_0 y una función $h : U \to V$ de clase C^p, con inversa $h^{-1} : V \to U$ de clase C^p, tales que*

$$f(h(x_1, \ldots, x_n)) = x_n.$$

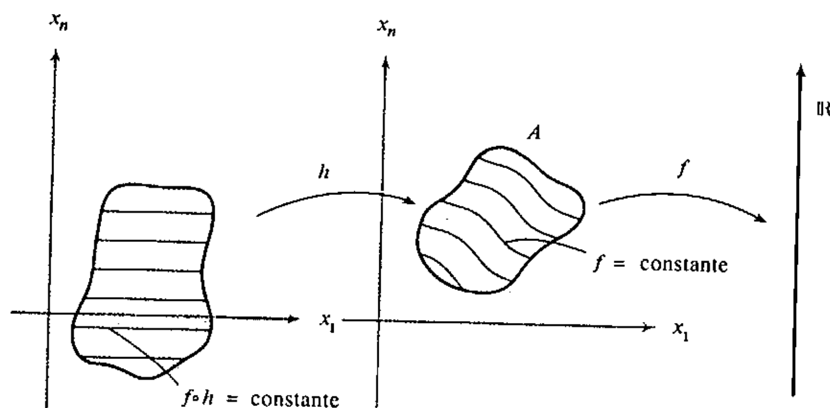


FIGURA 7.3-1 Un cambio de coordenadas puede rectificar los conjuntos de nivel de f

Este teorema tiene una generalización a las funciones $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$, dada en el ejercicio 18 al final del capítulo.

En la figura 7.3-1 se observa la plausibilidad del teorema. La misión de h es retorcer el espacio de tal modo que las superficies de nivel de f se conviertan en planos de dimensión $n - 1$. La condición $\mathbf{D}f(x_0) \neq 0$ se usa para garantizar que las superficies $f =$ constante son "no degeneradas" o, intuitivamente, que son suaves y de dimensión $n - 1$. Un ejemplo aclarará este punto.

7.3.2 Ejemplo *Sea* $f(x, y) = x^2 + y^2$. *¿Podemos "rectificar"* f *cerca de* $(0, 0)$? *¿Qué ocurre cerca de otros puntos?*

Solución La respuesta a la primera pregunta es "no, no necesariamente", pues $\mathbf{D}f(0, 0) = 0$. Esto está claro intuitivamente, pues las superficies con f constante se degeneran en $(0, 0)$, pasando de ser círculos a ser un punto (véase la figura 7.3-2). Es evidente que no hay forma de deformar las superficies $f =$ constante cerca de $(0, 0)$ en planos. Pero *sí podemos* hacerlo cerca de cualquier punto $(x_0, y_0) \neq (0, 0)$. ♦

7.3.3 Ejemplo *Sea* $f(x, y) = x^3 + x + y$. *¿Podemos "rectificar"* f *cerca de* $(0, 0)$?

Solución Sí, pues $\mathbf{D}f(0, 0) = (1, 1) \neq 0$. ♦

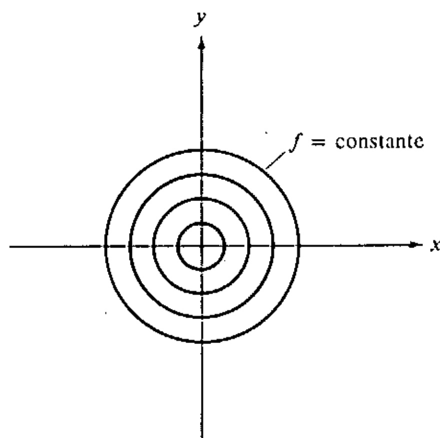


FIGURA 7.3-2 Las curvas de nivel de $x^2 + y^2$ se degeneran en (0, 0)

Ejercicios de §7.3

1. ¿En torno a qué puntos (x, y) podemos "rectificar" $f(x, y) = x^2 - y^2$?

2. Bosquéjense las gráficas de f = constante en el ejercicio 1 y explíquese la respuesta geométricamente.

3. ¿Se puede rectificar $f(x, y) = x^3 + y^2 + 1$ cerca de (0, 0)? ¿Y cerca de (0, 1)?

§7.4 Más consecuencias del teorema de la función implícita

El teorema de rectificación del dominio afirma que podemos encontrar una función h que "rectifica" el dominio de f de modo que $f \circ h$ sea simplemente una proyección. Por analogía con esto, podemos buscar una función g que "rectifique" la imagen de f de modo que $g \circ f$ sea como una "inserción".

7.4.1 Teorema de rectificación de la imagen *Sean $A \subset \mathbb{R}^p$ un conjunto abierto y $f : A \to \mathbb{R}^n$ una función de clase C^r, $r \geq 1$, y supóngase que $p \leq n$. Sea $x_0 \in A$ y supóngase que el rango de* $\mathbf{D}f(x_0)$ *es p. Entonces existen conjuntos abiertos U y V en $\mathbb{R}^n$ tales que $f(x_0) \in U$ y existe una función $g : U \to V$ de clase C^r con inversa $g^{-1} : V \to U$ también de clase C^r, tales que $g \circ f(x_1, \ldots, x_p) = (x_1, \ldots, x_p, 0, \ldots, 0)$ para todo $(x_1, \ldots, x_p) \in V$.*

Se ilustra este resultado intuitivamente en la figura 7.4-1, que debe compararse con la figura 7.3-1. En este caso, la función g aplana la *imagen* de f. Obsérvese que esto es intuitivamente correcto; esperamos que la imagen de f sea una "superficie" p-dimensional, de modo que debería poderse aplanar y convertirse en una porción de $\mathbb{R}^p$. Obsérvese que la imagen de una transformación lineal de rango p es un subespacio lineal de dimensión exactamente igual a p, de modo que este resultado expresa, en cierto sentido, una generalización del caso lineal.

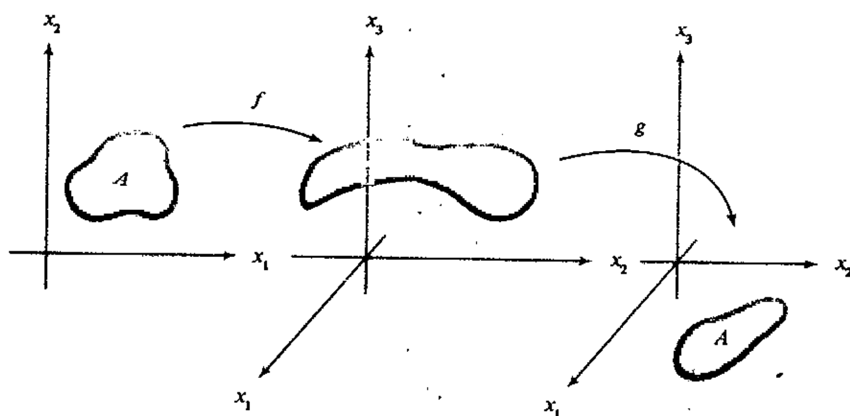


FIGURA 7.4-1 Un cambio de variables puede aplanar la imagen de f

Para usar los teoremas de rectificación, debemos tener que el rango de $\mathbf{D}f$ sea igual a la dimensión de su espacio imagen (o a la del espacio dominio). Sin embargo, podemos usar el teorema de la función inversa de nuevo, para concluir que si $\mathbf{D}f(x)$ tiene rango constante m en una vecindad de x_0, podemos rectificar el dominio de f con alguna función invertible h, de modo que $f \circ h$ sólo dependa de $x_1, \ldots, x_m$. Entonces también podemos rectificar la imagen. Ésta es la esencia del siguiente teorema y su corolario. *Grosso modo* dice que si $\mathbf{D}f$ tiene rango m en $\mathbb{R}^n$, entonces $n - m$ variables son redundantes y pueden eliminarse. Por ejemplo, si $f : \mathbb{R}^2 \to \mathbb{R}^1$, $f(x, y) = x - y$, $\mathbf{D}f$ tiene rango 1, por lo que podemos expresar f usando solamente una variable: sea $h(x, y) = (x + y, y)$, de modo que $f \circ h(x, y) = x$, que solamente depende de x.

Nota. Recuérdese que el ***rango*** de una transformación lineal es la dimensión de su imagen. De forma equivalente, el rango es el tamaño de la mayor submatriz cuadrada con determinante distinto de cero (consúltense los detalles en un texto de álgebra lineal).

7.4.2 Teorema del rango *Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^N$ (donde A es abierto en $\mathbb{R}^n$) una función C^r tal que $\mathbf{D}f(x)$ tenga rango m para todo x en una vecindad de $x_0 \in A$. Entonces existen un conjunto abierto $U \subset \mathbb{R}^n$, un conjunto abierto $V \subset \mathbb{R}^n$ con $x_0 \in V$ y una función $h : U \to V$ de clase C^r con inversa $h^{-1} : V \to U$ de clase C^r tal que $f \circ h$ depende solamente de $x_1, \ldots, x_m$. Es decir, $(f \circ h)(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n) = \tilde{f}(x_1, \ldots, x_m)$ para alguna función C^r, $\tilde{f}$. Véase la figura 7.4-2.*

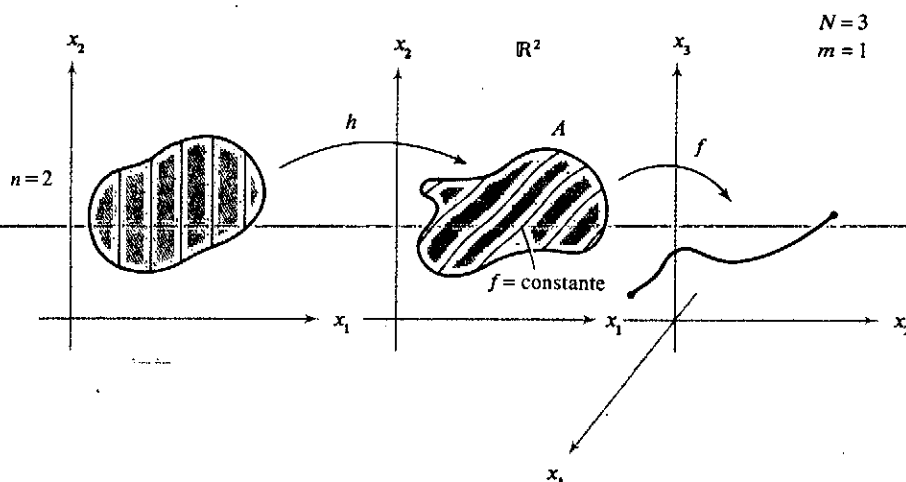


FIGURA 7.4-2 Un cambio de variables puede hacer que una transformación de rango m sólo dependa de m variables

7.4.3 Corolario *Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^N$ (donde A es abierto en $\mathbb{R}^n$) una función de clase C^r, $r \geq 1$, tal que $\mathbf{D}f(x)$ tenga rango m para todo x en una vecindad de $x_0 \in A$. Entonces existe un conjunto abierto $U_1 \subset \mathbb{R}^n$, un conjunto abierto $U_2 \subset \mathbb{R}^n$ tal que $x_0 \in U_2$, un conjunto abierto V_1 en torno a $f(x_0)$, un conjunto abierto $V_2 \subset \mathbb{R}^N$, y funciones $h : U_1 \to U_2$ y $g : V_1 \to V_2$ de clase C^r con inversas de clase C^r tales que $(g \circ f \circ h)(x_1, \ldots, x_n) = (x_1, \ldots, x_m, 0, \ldots, 0)$.*

En **§7.7** damos más aplicaciones del teorema de la función implícita a la teoría de superficies y a los multiplicadores de Lagrange (problemas de extremos condiciona-

dos). Además, en las secciones restantes del capítulo se estudian algunos temas (opcionales) mediante el mismo método o uno similar.

7.4.4 Ejemplo *Sea* $f : \mathbb{R}^2 \to \mathbb{R}^3$, $(x, y) \mapsto (x + y^3, xy, y + y^2)$. *¿Se puede "rectificar" la imagen cerca de* $(0, 0)$*?*

Solución En este caso usamos el teorema de rectificación de la imagen. En primer lugar, calculamos la matriz jacobiana:

$$\begin{pmatrix} 1 & 3y^2 \\ y & x \\ 0 & 1 + 2y \end{pmatrix},$$

que en $(0, 0)$ es

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Esta matriz tiene rango 2 (pues existe una submatriz 2×2 con determinante distinto de cero). Por lo tanto, el teorema de rectificación de la imagen se puede aplicar, por lo que podemos rectificar la imagen. [Intuitivamente, será una superficie bidimensional cerca de $(0, 0)$.] ◆

7.4.5 Ejemplo *Sea* $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 + y$. *¿Se puede expresar f como función de una sola variable cerca de* $(0, 0)$*?*

Solución Sí, pues (por el teorema del rango) $\mathbf{D}f(0, 0) = (0, 1) \neq 0$; por continuidad, $\mathbf{D}f$ también es distinto de cero cerca de $(0, 0)$. Así, $\mathbf{D}f$ tiene rango 1 cerca de $(0, 0)$. Obsérvese que también se puede responder usando el teorema de rectificación del dominio. ◆

Ejercicios de §7.4

1. Sea $f : \mathbb{R}^2 \to \mathbb{R}^3$, $(x, y) \mapsto (x + y^2, xy, y^2)$. ¿Puede rectificarse la imagen cerca de $(0, 0)$? ¿Y cerca de $(0, 1)$?

2. ¿Qué dice el teorema del rango acerca de la función $f : \mathbb{R}^3 \to \mathbb{R}^2$ definida por $(x, y, z) \mapsto (x^2 + 2y^2, z^2 + 3xy)$ cerca de $(0, 0, 0)$? ¿Y cerca de $(0, 1, 0)$?

3. ¿Qué dice el corolario **7.4.3** acerca de $f : \mathbb{R}^3 \to \mathbb{R}^3$, $(x, y, z) \mapsto (x + 2y, 6x + 12y, x + y^3 + z^3)$ cerca de $(0, 0, 0)$?

4. Examínese la afirmación del corolario **7.4.3** en el caso donde f sea una transformación lineal.

§7.5 Un teorema de existencia para ecuaciones diferenciales ordinarias

En **§5.7** usamos el principio de la aplicación contractiva para iniciar nuestro estudio de las ecuaciones diferenciales e integrales. Ahora estudiaremos las ecuaciones diferenciales con un poco más de profundidad. En cálculo aprendimos a resolver las ecuaciones diferenciales simples de forma explícita; por ejemplo, la solución $d^2x/dt^2 + k^2x = 0$ es $x(t) = A\cos(kt - \omega)$ para A y ω constantes. Es interesante investigar si una ecuación diferencial general tiene o no solución. Los métodos que analizaremos son constructivos y adecuados para el cálculo numérico; es decir, se construye una sucesión específica de soluciones aproximadas; un ejemplo aclarará estas cuestiones.

7.5.1 Ejemplo *Considérese la ecuación no lineal $dx/dt = x^2$, $x(0) = 1$. ¿Podemos calcular $x(1)$?*

Solución En este caso, podemos resolver la ecuación de forma explícita, mediante la separación de variables. Escribimos $dx/x^2 = dt$ e integramos, obteniendo $-1/x = t + C$; es decir, $x = -1/(t + C)$. En $t = 0$, $x = 1$, por lo que $C = -1$. Así, $x = 1/(1 - t)$ es nuestra solución. Está claro que ésta es la única solución que empieza en $t = 0$ con $x(0) = 1$. En $t = 1$, la solución $x(t)$ "explota", por lo que $x(1)$ *no está definido*. Así, no podemos determinar una solución diferenciable $x(t)$ definida para todo $t \geq 0$ (*f*igura 7.5-1). ♦

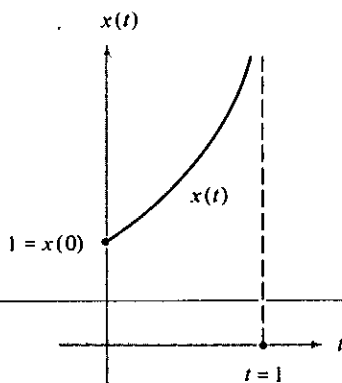


FIGURA 7.5-1 La solución de $x'(t) = x^2$ que pasa por $(0,1)$ "explota" cuando t tiende a 1

Este ejemplo muestra el importante hecho de que, en general, *las soluciones $x(t)$ pueden estar definidas y ser derivables sólo en un pequeño intervalo de t.*

Es importante también hacer otra afirmación. Si permitimos ecuaciones diferenciales vectoriales, entonces las ecuaciones de orden superior se pueden reducir a ecuaciones de primer orden.

7.5.2 Ejemplo *Escríbase $d^2x/dt^2 + kx = 0$ como una ecuación de primer orden.*

Solución Sea $y = dx/dt$ y escríbase

$$\begin{cases} dx/dt = y \\ dy/dt = -kx, \end{cases}$$

que es de primer orden en el vector (x, y) y es equivalente a la ecuación original. ♦

Ahora daremos el principal teorema de existencia y unicidad, donde denotaremos $B(x_0, r)$ la bola *cerrada* de radio r y centro x_0, de modo que

$$B(x_0, r) = \{y \in \mathbb{R}^n \mid \|x_0 - y\| \leq r\}.$$

7.5.3 Teorema *Sea $f : [-a, a] \times B(x_0, r) \to \mathbb{R}^n$ una aplicación continua. Sea $C = \sup\{\|f(t, x)\| \mid (t, x) \in [-a, a] \times B(x_0, r)\}$. Supóngase que existe K constante tal que*

$$\|f(t,x) - f(t,y)\| \leq K\|x - y\| \qquad \text{(L)}$$

para todo $t \in [-a, a], x, y \in B(x_0, r)$. Sea $b < \text{mín}\{a, r/C, 1/K\}$. Entonces existe una única transformación continuamente diferenciable $x : [-b, b] \to B(x_0, r) \subset \mathbb{R}^n$ tal que

$$x(0) = x_0 \quad (\textit{condición inicial}) \quad y \quad \frac{dx}{dt} = f(t, x(t)).$$

La condición principal sobre f es la ***condición de Lipschitz*** (L) enunciada en el teorema, donde el número K es la ***constante de Lipschitz***. Si f cumple la condición, decimos que es ***Lipschitz*** con respecto de la variable x. Para verificar esta condición, con frecuencia utilizamos las ideas de los capítulos 4, 5 y 6: si $\mathbf{D}_x f(t, x)$ denota la derivada de f con t fijo y $\|\mathbf{D}_x f(t, x)y\| \leq K\|y\|$ para todo $y \in \mathbb{R}^n$, entonces f es Lipschitz, con constante de Lipschitz K. Por ejemplo, si $n = 1$, esto se cumple si $|\partial f(t, x)/\partial x| \leq K$ para $-a \leq t \leq a, -r \leq x - x_0 \leq r$. Podemos ver esto si usamos la regla de la cadena como sigue:

$$\frac{d}{ds} f(t, y + s(x - y)) = \mathbf{D}_x f(t, y + s(x - y)) \cdot (x - y),$$

de modo que al integrar entre $s = 0$ y $s = 1$, tenemos

$$f(t,x) - f(t,y) = \int_0^1 \mathbf{D}_x f(t, y + s(x - y)) \cdot (x - y)\, ds.$$

Se obtiene el resultado al usar los valores absolutos. Obsérvese que si f es C^1, dicha K siempre existe (¿por qué?).

Con frecuencia, f es independiente de t, en cuyo caso decimos que tenemos un ***sistema autónomo.*** Si f sólo es continua, se cumple la existencia (pero no la unicidad) de $x(t)$; véase el ejercicio 45 al final del capítulo 5.

La idea de la demostración de **7.5.3** es usar aproximaciones sucesivas; se comienza con

$$x_1(t) = x_0,$$

y se escribe

$$x_2(t) = x_0 + \int_0^t f(s, x_1(s))\, ds$$

$$x_3(t) = x_0 + \int_0^t f(s, x_2(s))\, ds$$

$$\vdots$$

$$x_n(t) = x_0 + \int_0^t f(s, x_{n-1}(s))\, ds.$$

Después se demuestra que $x_n(t)$ converge a una función $x(t)$ que satisface

$$x(t) = x_0 + \int_0^t f(s, x(s))\, ds$$

Como $f(t)$ y $x(t)$ son continuas, el teorema fundamental del cálculo afirma que el miembro derecho de la última ecuación es una función derivable de t, con derivada $f(t, x(t))$. Así, el miembro izquierdo $x(t)$ debe ser derivable y $x'(t) = f(t, x(t))$. Como la integral de 0 a 0 se anula, está claro que $x(0) = x_0$. Si comparamos esto con §**5.7** vemos que buscamos un punto fijo para la transformación de una función en otra dada por

$$(Ty)(t) = x_0 + \int_0^t f(s, y(s))\, ds.$$

El proceso de iteración que hemos bosquejado es el ***proceso de iteración de Picard*** y es igual al que se usa en forma abstracta para la demostración del principio de la aplicación contractiva. La condición de Lipschitz y la elección del intervalo $[-b, b]$ se usan para mostrar que T es una contracción en el espacio $C([-b, b])$.

7.5.4 Ejemplo *¿Qué valores de b produce el teorema* **7.5.3** *para el problema del ejemplo* **7.5.1***?*

Solución En este caso, $dx/dt = x^2$ y $x(0) = 1$. Por el momento, sean a, r indeterminadas. Ahora,

$$\begin{aligned} C &= \sup\{|f(t,x)| \mid -a \le t \le a, -r \le x - 1 \le r\} \\ &= \sup\{x^2 \mid -r \le x - 1 \le r\} = (r+1)^2 \end{aligned}$$

(véase la figura 7.5-2). Así, $r/C = r/(r+1)^2$. Además, $\partial f/\partial x = 2x$ y entonces

$$K = \sup\{2|x| \mid -r \le x - 1 \le r\} = 2(r+1).$$

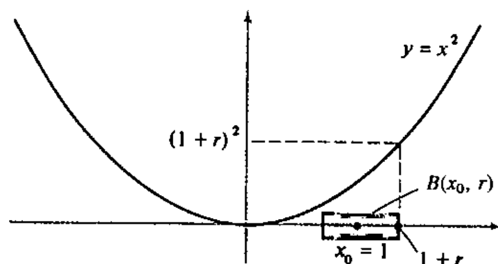


FIGURA 7.5-2 Aplicación del teorema **7.5.3** al caso $f(x, t) = x^2$

Como a no está relacionada con las estimaciones de C y K, elegimos a lo más grande posible, de modo que no interfiera; digamos que $a = 100$. Por el teorema, elegimos

$$b < \min\left\{\frac{r}{(r+1)^2}, \quad \frac{1}{2(r+1)}\right\}.$$

Esto sirve para cualquier $r > 0$. Ambos valores son iguales a $\frac{1}{4}$ si $r = 1$. Si $0 < r < 1$, el primero es menor y crece de 0, en $r = 0$, a $\frac{1}{4}$, en $r = 1$. Si $r > 1$, el segundo es menor y decrece desde $\frac{1}{4}$ cuando r crece a partir de 1. Así, lo mejor que podemos obtener de esto es hacer $r = 1$ y concluir que el tiempo de existencia es al menos tan grande como cualquier b < $\frac{1}{4}$. Esto no es una buena estimación, como se ve si analizamos el problema directamente (con un tiempo de existencia < 1), pero podemos volver a aplicar el teorema para obtener un nuevo tiempo de existencia en $t = \frac{1}{4}$ y extender nuestro trabajo a cualquier $t < 1$. Pero nunca podremos pasar de $t = 1$. ♦

Ejercicios de §7.5

1. Resuélvase $dx/dt = 1 + x^2$, $x(0) = 0$ por el método de aproximaciones sucesivas. ¿Está definida $x(t)$ para todo $t \geq 0$?

2. Úsese el teorema **7.5.3** para calcular b en el ejercicio 1.

3. Muéstrese que $dx/dt = \sqrt{x}$, $x(0)$ tiene dos soluciones:

$$x(t) = 0 \text{ para todo } t \quad \text{y} \quad x(t) = \begin{cases} 0, & t \leq 0 \\ t^2/4, & t > 0. \end{cases}$$

¿Contradice esto el teorema **7.5.3**?

4. Considérese la ecuación $dx/dt = te^{x^2}$ sen x, $x(0) = 1$. Obtenga una estimación del intervalo donde se puede definir la solución $x(t)$.

5. Sea A una matriz $n \times n$ y considérese el sistema lineal

$$\frac{dx}{dt} = A \cdot x(t); \quad x(t) \in \mathbb{R}^n.$$

a. Muéstrese que una solución es

$$x(t) = e^{tA}x(0), \qquad \text{donde } e^B = \sum_{n=0}^{\infty} \frac{B^n}{n!}.$$

b. El tiempo de existencia en este caso se extiende para todo t. ¿Se podría obtener esta conclusión también del teorema **7.5.3**?

§7.6 Lema de Morse

En el capítulo 6 vimos que la hessiana de una función $f : \mathbb{R}^n \to \mathbb{R}$ en un punto crítico determinaba el comportamiento local de f cerca de ese punto. El lema de Morse lleva este análisis más lejos. Por ejemplo, establece que si f tiene un mínimo local no degenerado en x_0, entonces f no sólo *parece* un paraboloide, sino que podemos cambiar de coordenadas de modo que f *realmente "sea" un paraboloide* en las nuevas coordenadas. El "lema" (un teorema, en realidad) también se aplica a las superficies de silla.

El lema de Morse es fundamental para el trabajo avanzado en topología y análisis, pero incluso aquí nos ayuda a comprender mejor la forma de las funciones cerca de un punto crítico.

7.6.1 Lema de Morse *Sean $A \subset \mathbb{R}^n$ un conjunto abierto y $f: A \to \mathbb{R}$ una función suave (infinitamente derivable). Supóngase que $\mathbf{D}f(x_0) = 0$ y que la hessiana de f en x_0 no es singular. Entonces, existen una vecindad U de x_0 y una vecindad V de 0 en $\mathbb{R}^n$, así como una aplicación suave $g: V \to U$ con inversa suave tales que $f \circ g = h$ tiene la forma*

$$h(y) = f(x_0) - [y_1^2 + y_2^2 + \cdots + y_\lambda^2] + [y_{\lambda+1}^2 + \cdots + y_n^2],$$

donde λ es algún entero fijo entre 0 y n.

Decimos que g es un ***cambio de coordenadas*** y hablamos de $y = g^{-1}(x)$ como las ***nuevas coordenadas***.

Un punto crítico en el que la matriz hessiana $\Delta = \partial^2 f/\partial x_i \partial x_j$ no es singular es un ***punto crítico no degenerado***. Así, el lema de Morse da una descripción bastante completa de las funciones en la vecindad de un punto crítico no degenerado. El número λ se denomina ***índice*** del punto crítico. La figura 7.6-1 muestra las gráficas de las formas cuadráticas $-y_1^2 - \cdots - y_\lambda^2 + y_{\lambda+1}^2 + \cdots + y_n^2$ para distintos índices en $\mathbb{R}^2$.

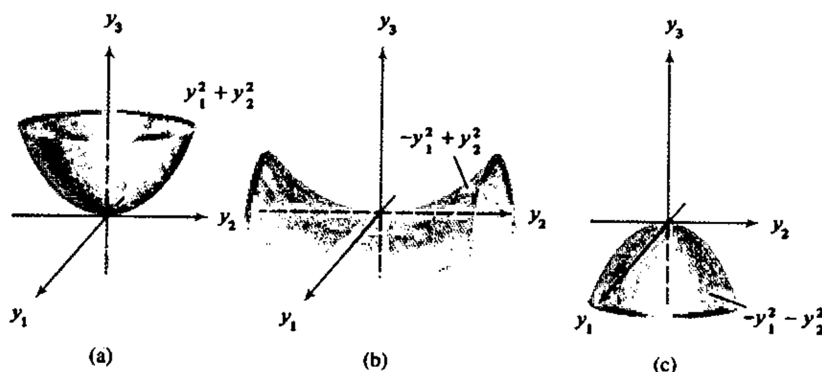


FIGURA 7.6-1 (a) Índice = 0; (b) índice = 1; (c) índice = 2

Para las funciones de dos variables es fácil determinar el índice: si Δ es definida negativa, el índice es 2; si Δ es definida positiva, el índice es cero, y en cualquier otro caso, es 1. En general, para determinar el índice hay que saber un poco más de álgebra lineal. El lector con tal conocimiento puede verificar que el índice es exactamente el número de autovalores negativos de Δ.

7.6.2 Ejemplo *¿Cuál es la forma de la superficie $z = x^2 + 2xy + 2y^2 + y^3$ cerca de $(0, 0)$?*

Solución Tenemos un punto crítico en (0, 0), y la hessiana es

$$\Delta = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right] = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix},$$

que es definida positiva, pues $2 > 0$ y $\det(\Delta) = (2)(4) - (2)(2) > 0$. Así, el índice es 0 y cerca de (0, 0) la superficie es aproximadamente un paraboloide y, en algún otro sistema de coordenadas, es exactamente un paraboloide. ♦

7.6.3 Ejemplo *Calcúlese el índice de* $x^2 - 3xy + y^2 + 8xy^2 + 6$ *en* (0, 0).

Solución (0, 0) es un punto crítico y la hessiana es

$$\Delta = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}.$$

Como $2 > 0$ y $\det \Delta = -5 < 0$, esta matriz no es definida positiva ni definida negativa. Así, tenemos un índice 1 y un punto silla en (0, 0). ♦

Ejercicios de §7.6

1. Calcúlese el índice de $2x^2 + 6xy - y^2 + y^4$ en (0, 0).

2. ¿Cuál es la forma de la superficie $x^2 + 3xy - y^2$ en (0, 0)?

3. ¿Se aplica el lema de Morse a $x^2 - 2xy + y^2$?

4. Sea $f(x, y) = x^2 + y^2 + 3y^3 + 8x^4 + x^2e^x \operatorname{sen} x + 6$. Muéstrese que existen nuevas coordenadas ξ, η, donde

$$\xi = \xi(x, y), \quad \eta = \eta(x, y),$$

para las cuales

$$f(x, y) = \xi^2 + \eta^2 + 6$$

en una vecindad de (0, 0).

5. **a.** Si f tiene un punto crítico no degenerado en $x_0 \in \mathbb{R}^n$, muéstrese que existe una vecindad de x_0 que no contiene otros puntos críticos.

   b. ¿Cómo son los puntos críticos de la función $f(x, y) = x^2y^2$?

§7.7 Extremos condicionados y multiplicadores de Lagrange

En algunos problemas se pide maximizar una función sujeta a ciertas *restricciones*. Situaciones como ésta se presentan, por ejemplo, en economía: supóngase que estamos vendiendo dos tipos de artículos, digamos, I y II, y sean x, y la cantidad vendida de cada uno. Sea $f(x, y)$ la ganancia obtenida al vender x artículos del tipo I e y del tipo II. Como nuestra producción está limitada por nuestro capital, estamos sujetos a trabajar con una relación, digamos, $g(x, y) = 0$. Así, queremos maximizar $f(x, y)$ entre los x, y que satisfagan $g(x, y) = 0$. La condición $g(x, y) = 0$ se denomina ***restricción*** del problema. El propósito de esta sección es analizar algunos métodos que nos permitan manejar ésta y otras situaciones similares.

7.7.1 Teorema del multiplicador de Lagrange *Sean $f : U \subset \mathbb{R}^n \to \mathbb{R}$ y $g : U \subset \mathbb{R}^n \to \mathbb{R}$ funciones C^1 dadas. Sean $x_0 \in U$, $g(x_0) = c_0$, y $S = g^{-1}(c_0)$, el conjunto de nivel de g con valor c_0. Supóngase que $\nabla g(x_0) \neq 0$. Si $f \mid S$, lo que denota la restricción de f a S (es decir, a aquellos $x \in U$ que satisfagan $g(x) = c_0$), tiene un máximo o un mínimo en x_0, entonces existe un número real λ tal que*

$$\nabla f(x_0) = \lambda \nabla g(x_0).$$

La *idea* de la demostración es la siguiente. Recuérdese que el espacio tangente a S en x_0 se define como el espacio ortogonal a $\nabla g(x_0)$ (véase **§6.6**). Motivamos esta definición considerando las tangentes a los arcos $c(t)$ que están en S, como sigue: si $c(t)$ es un arco en S con $c(0) = x_0$, entonces $c'(0)$ es un vector tangente a S en x_0, pues

$$\frac{d}{dt} g(c(t)) = \frac{d}{dt} c_0 = 0,$$

y, por otro lado, por la regla de la cadena,

$$\frac{d}{dt} g(c(t))\big|_{t=0} = \nabla g(x_0) \cdot c'(0),$$

de modo que $c'(0)$ es ortogonal a $\nabla g(x_0)$.

Si $f \mid S$ tiene un máximo en x_0, entonces ciertamente $f(c(t))$ tiene un máximo en $t = 0$. Por lo tanto,

$$0 = \frac{d}{dt} f(c(t))\big|_{t=0} = \nabla f(x_0) \cdot c'(0).$$

Así, $\nabla f(x_0)$ es también perpendicular al espacio tangente a S en x_0, por lo que $\nabla f(x_0)$ y $\nabla g(x_0)$ son paralelos. Como $\nabla g(x_0) \neq 0$, esto implica que $\nabla f(x_0)$ es un múltiplo de $\nabla g(x_0)$, que es precisamente la conclusión del teorema.

El siguiente corolario extrae algunas consecuencias geométricas de esta demostración.

7.7.2 Corolario *Si f, al estar restringida a una superficie S, tiene un máximo o un mínimo en x_0, entonces $\nabla f(x_0)$ es perpendicular a S en x_0 (véase la figura 7.7-1).*

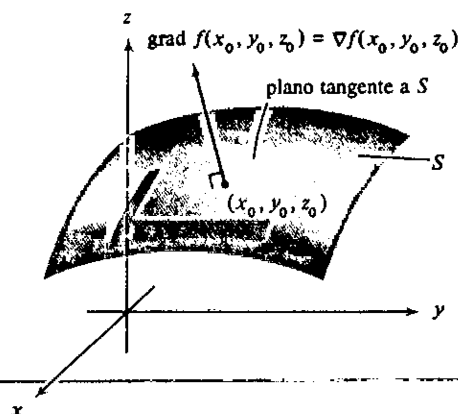


FIGURA 7.7-1 La geometría de los extremos condicionados

Estos resultados nos dicen que para determinar los extremos con restricciones de f, debemos buscar entre los puntos x_0 que satisfagan las conclusiones del teorema o del corolario. Daremos algunos ejemplos del uso de cada uno. Cuando se usa el método del teorema **7.7.1**, debemos buscar un punto x_0 y una constante λ, llamada ***multiplicador de Lagrange***, de modo que $\nabla f(x_0) = \lambda \nabla g(x_0)$. Este método tiene una naturaleza más analítica, mientras que el método del corolario **7.7.2** es más geométrico. En cualquier caso, las técnicas de multiplicadores de Lagrange son más eficaces cuando se usan conjuntamente con la intuición (no en sustitución de ésta) obtenida en el estudio del cálculo elemental.

Un *criterio de la segunda derivada* para los extremos condicionados se puede ver en *Cálculo vectorial* de Marsden y Tromba (*op. cit.*).

Ejemplo 7.7.3 *Sea $S \subset \mathbb{R}^2$ una recta que pasa por $(-1, 0)$ con una inclinación de 45° y sea $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, t) \mapsto x^2 + t^2$. Determínese el mínimo de f en S.*

Solución En este caso, $S = \{(x, t) \mid t - x - 1 = 0\}$, de modo que elegimos $g(x, t) = t - x - 1$. El extremo de $f \mid S$ debe ser buscado de entre los puntos donde ∇f es ortogonal a S; es decir, con una inclinación de $-45°$. Pero $\nabla f(x, t) = (2x, 2t)$ y tiene la pendiente de-

seada cuando $x = -t$; es decir, (x, t) está en la recta L que pasa por el origen, con una inclinación de $-45°$. Esto puede ocurrir en un punto (x, t) que está en el conjunto S sólo para el único punto que está en la intersección de L y S (véase la figura 7.7-2). La referencia a las curvas de nivel de f indica que este punto, $(-1/2, 1/2)$, es un mínimo relativo de $f \mid S$ (pero no de f). ♦

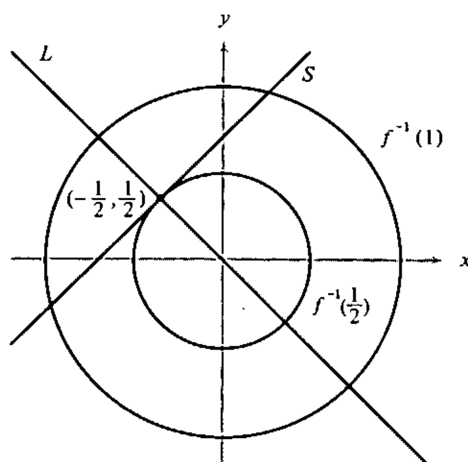


FIGURA 7.7-2 Localización de los puntos críticos de la restricción de f a S

7.7.4 Ejemplo *Sea $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $(x, y) \mapsto x^2 - y^2$, y sea S la circunferencia de radio* 1 *con centro en el origen. Determínense los puntos críticos de f restringida a S.*

Solución En este caso, $S = g^{-1}(1)$, donde $g : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x^2 + y^2$. Las curvas de nivel, los espacios tangentes y los gradientes aparecen en la figura 7.7-3. Está claro que el gradiente de f es ortogonal a S en los cuatro puntos $(0, \pm 1)$, $(\pm 1, 0)$, que son máximos y mínimos relativos, respectivamente, de $f \mid S$.

Este problema se puede resolver analíticamente mediante el método de los multiplicadores de Lagrange. Está claro que

$$\nabla f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = (2x, -2y)$$

y

$$\nabla g(x, y) = (2x, 2y).$$

Así, buscamos puntos (x, y) con $x^2 + y^2 = 1$ para los que exista λ tal que $(2x, -2y) = \lambda(2x, 2y)$. Esto nos da tres ecuaciones: $2x = \lambda \cdot 2x$, $-2y = \lambda \cdot 2y$, y $x^2 + y^2 = 1$, de las que podemos obtener los valores de las tres incógnitas x, y y λ. De $2x = \lambda 2x$, podemos concluir que o bien $x = 0$ o bien $\lambda = 1$. Si $x = 0$, entonces $y = \pm 1$ y $-2y = \lambda 2y$ implica $\lambda = -1$. Si $\lambda = 1$, entonces $y = 0$ y $x = \pm 1$. Así, obtenemos los mismos puntos $(0, \pm 1)$, $(\pm 1, 0)$ que antes. Como ya mencionamos, el método solamente localiza los posibles extremos; el hecho de que sean máximos, mínimos o puntos silla debe determinarse mediante inspección o por otros medios. ♦

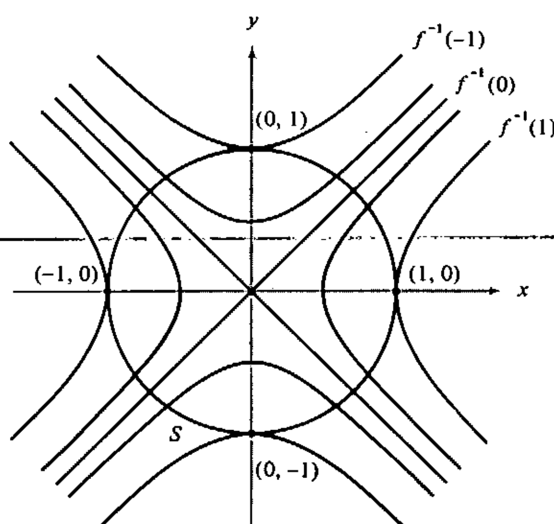


FIGURA 7.7-3 Las hipérbolas son tangentes a S en los puntos críticos de $f|S$

Si la superficie S está definida por varias restricciones,

$$\begin{aligned} g_1(x_1, \ldots, x_n) &= c_1 \\ g_2(x_1, \ldots, x_n) &= c_2 \\ &\vdots \\ g_k(x_1, \ldots, x_n) &= c_k \end{aligned}$$

(hasta ahora sólo teníamos una g), entonces el teorema **7.7.1** se puede generalizar como sigue: si f tiene un máximo o un mínimo en x_0 sobre S y $\nabla g_1(x_0), \ldots, \nabla g_k(x_0)$ son

linealmente independientes, entonces existen constantes $\lambda_1, \ldots, \lambda_k$ tales que

$$\nabla f(x_0) = \lambda_1 \nabla g_1(x_0) + \cdots + \lambda_k \nabla g_k(x_0).$$

Esto se puede demostrar generalizando el método utilizado para demostrar el teorema **7.7.1** y se deja para el lector interesado. Ahora daremos un ejemplo del uso de esta formulación general.

7.7.5 Ejemplo *Determínense los puntos extremos de* $f(x, y, z) = x + y + z$ *sujeta a las condiciones* $x^2 + y^2 = 2$ *y* $x + z = 1$.

Solución En este caso tenemos dos restricciones,

$$g_1(x, y, z) = x^2 + y^2 - 2 = 0$$

y

$$g_2(x, y, z) = x + z - 1 = 0.$$

Así, debemos determinar x, y, z y λ_1 y λ_2 tales que

$$\nabla f(x, y, z) = \lambda_1 \nabla g_1(x, y, z) + \lambda_2 \nabla g_2(x, y, z)$$

y

$$g_1(x, y, z) = 0$$
$$g_2(x, y, z) = 0;$$

es decir, debemos resolver

$$1 = \lambda_1 \cdot 2x + \lambda_2 \cdot 1$$
$$1 = \lambda_1 \cdot 2y + \lambda_2 \cdot 0$$
$$1 = \lambda_1 \cdot 0 + \lambda_2 \cdot 1$$

y

$$x^2 + y^2 = 2$$
$$x + z = 1.$$

Éstas son cinco ecuaciones para x, y, z, λ_1 y λ_2. De la tercera, $\lambda_2 = 1$, y entonces $2x\lambda_1 = 0$, $2y\lambda_1 = 1$. Como la segunda implica que $\lambda_1 \neq 0$, tenemos que $x = 0$. Así, $y = \pm\sqrt{2}$ y $z = 1$. Por lo tanto, nuestros puntos son $(0, \pm\sqrt{2}, 1)$. Las restricciones $x^2 + y^2 = 2$ y $x + z = 1$ definen juntas un conjunto compacto en $\mathbb{R}^3$, la elipse que se forma al intersecar un cilindro y un plano. La función continua f debe alcanzar un máximo y un mínimo en él, y esto debe ocurrir para algunos de los puntos que hemos determinado. En consecuencia, $f(0, \sqrt{2}, 1) = 1 + \sqrt{2}$ debe ser el máximo y $f(0, -\sqrt{2}, 1) = 1 - \sqrt{2}$ el mínimo. ♦

7.7.6 Ejemplo *Maximícese $f(x, y, z) = x + z$ sujeta a la restricción $x^2 + y^3 + z^2 = 1$.*

Solución Aquí usamos el teorema 7.7.1. Buscamos λ y (x, y, z) tales que

$$
\begin{aligned}
1 &= 2x\lambda \\
0 &= 2y\lambda \\
1 &= 2z\lambda
\end{aligned}
$$

y

$$x^2 + y^2 + z^2 = 1.$$

De la primera ecuación, $\lambda \neq 0$, por lo que la segunda da $y = 0$ y la primera y tercera dan $z = x$. La cuarta se convierte en $x^2 + x^2 = 1$, de modo que $x = \pm 1/\sqrt{2}$. Nuestros puntos son $(1/\sqrt{2}, 0, 1/\sqrt{2})$ y $(-1/\sqrt{2}, 0, -1/\sqrt{2})$. La restricción define un conjunto compacto (la esfera unidad en $\mathbb{R}^3$) donde la función continua f debe alcanzar un máximo y un mínimo. Éstos deben estar entre nuestros puntos, por lo que comparamos los valores: $f(1/\sqrt{2}, 0, 1/\sqrt{2}) = \sqrt{2}$ debe ser el máximo y $f(-1/\sqrt{2}, 0, -1/\sqrt{2}) = -\sqrt{2}$ debe ser el mínimo. ♦

7.7.7 Ejemplo *Determínese el máximo volumen de una caja rectangular, si está sujeta a la restricción de que el área de su superficie está fija en 10 metros cuadrados.*

Solución En este caso, si x, y, z son las longitudes de los lados, el volumen es $f(x, y, z) = xyz$. La restricción es $2(xy + xz + yz) = 10$, es decir, $xy + xz + yz = 5$. Así, nuestras condiciones son

$$
\begin{aligned}
yz &= \lambda(y + z). \\
xz &= \lambda(x + z). \\
xy &= \lambda(x + y). \\
xy + xz + yz &= 5.
\end{aligned}
$$

En primer lugar, $x \neq 0$, ya que $x = 0$ implica $yz = 5$ y $0 = \lambda z$, de modo que $\lambda = 0$ y $yz = 0$. Análogamente, $y \neq 0$, $z \neq 0$, y $x + y \neq 0$, etcétera. La eliminación de λ en las primeras dos ecuaciones da $yz/(y + z) = xz/(x + z)$, lo que a su vez indica que $x = y$. De forma similar, $y = z$. Usando la última ecuación, $3x^2 = 5$, de modo que $x = \sqrt{5/3}$. Así, $x = y = z = \sqrt{5/3}$, de modo que $xyz = (5/3)^{3/2}$. Ésta es la solución. Debería estar claro geométricamente que el máximo ocurre cuando $x = y = z$. ♦

Ejercicios de §7.7

En los ejercicios 1 al 5, determínense los extremos de f sujeta a las restricciones dadas.

1. $f(x,y,z) = x - y + z,\ x^2 + y^2 + z^2 = 2.$

2. $f(x,y) = x - y,\ x^2 - y^2 = 2.$

3. $f(x,y) = x,\ x^2 + 2y^2 = 3.$

4. $f(x,y,z) = x + y + z,\ x^2 - y^2 = 1,\ 2x + z = 1.$

5. $f(x,y) = 3x + 2y,\ 2x^2 + 3y^2 = 3.$

6. La Corporación Supranacional del Lodo produce lodo mediante un equipo y un material con un costo de p = \$243 por unidad y una mano de obra con un salario de w = \$16 por hora. Si se usan x unidades de equipo/material e y horas de trabajo, entonces se producen $20x^{3/4}y^{1/4}$ litros de lodo. Si la compañía tiene un presupuesto de B = \$51,840,000, determínese la máxima cantidad de lodo que puede producir y la cantidad de equipo/material y de horas de trabajo necesarias para producirlo.

Demostraciones de los teoremas del capítulo 7

7.1.1 Teorema de la función inversa *Sean $A \subset \mathbb{R}^n$ un conjunto abierto y $f : A \subset \mathbb{R}^n \to \mathbb{R}^n$ de clase C^1 (es decir, supóngase que* **D***f existe y es continua). Sea $x_0 \in A$ y supóngase que $Jf(x_0) \neq 0$. Entonces, existen una vecindad U de x_0 en A y una vecindad abierta W de $f(x_0)$, tal que $f(U) = W$ y la restricción de f a U tiene una inversa C^1, $f^{-1} : W \to U$. Además, para $y \in W$ y $x = f^{-1}(y)$, tenemos*

$$\mathbf{D}f^{-1}(y) = [\mathbf{D}f(x)]^{-1},$$

la inversa de **D***$f(x)$, lo que significa su inversa como transformación lineal (correspondiente a la matriz inversa). Si f es de clase C^p, $p \geq 1$, también lo es f^{-1}.*

Demostración La demostración del teorema de la función inversa no es particularmente fácil en lo tocante a sus detalles técnicos, pero este teorema representa una piedra angular del análisis y por lo tanto debería dominarse. Vamos a hacer la demostración en $\mathbb{R}^n$, pero en realidad el teorema es válido, con ideas similares, en un espacio de Banach (véanse los detalles en Abraham, Marsden y Ratiu, *Manifolds, Tensor Analysis and Applications*, Springer-Verlag, Nueva York, 1988). La demostración se basa en el principio de la aplicación contractiva; esta técnica es útil, pues se aplica en muchas situaciones. Usaremos el siguiente caso particular.

Lema 1 *Sea M un subconjunto cerrado de $\mathbb{R}^n$ y d la función distancia en $\mathbb{R}^n$. Sea f una aplicación de M en M. Supóngase que existe K constante, donde $0 < K < 1$, tal que para cualesquiera dos puntos x e y en M, tenemos $d(f(x), f(y)) < Kd(x, y)$. Entonces existe un único $x_* \in M$ tal que $f(x_*) = x_*$ (x_* es un* ***punto fijo*** *de f).*

Antes de iniciar la demostración del teorema de la función inversa, es útil disponer de un lema técnico acerca del conjunto de transformaciones lineales invertibles (o, de forma equivalente, el conjunto de matrices invertibles). Una matriz $m \times n$ (o una transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$) es simplemente una mn-upla de números reales, ya que una matriz A con entradas (a_{ij}) se puede considerar como una mn-upla $(a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{m1}, \ldots, a_{mn})$. Así, tiene sentido decir que cierto subconjunto del conjunto de todas las matrices es abierto o que una transformación del conjunto de las matrices $m \times n$ al conjunto de las matrices $p \times q$ es diferenciable. Sea $L(\mathbb{R}^n, \mathbb{R}^n)$ el conjunto de las matrices $n \times n$ (o de las transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^n$) y $GL(\mathbb{R}^n, \mathbb{R}^n)$ el conjunto de todas las matrices invertibles (o transformaciones lineales invertibles de $\mathbb{R}^n$ en $\mathbb{R}^n$), llamado grupo general lineal. Así, $GL(\mathbb{R}^n, \mathbb{R}^n) = \{A \in L(\mathbb{R}^n, \mathbb{R}^n) \mid \det A \neq 0\}$. Sea $\mathcal{L}^{-1} : GL(\mathbb{R}^n, \mathbb{R}^n) \to GL(\mathbb{R}^n, \mathbb{R}^n)$ la aplicación que lleva una matriz invertible A en su inversa A^{-1}.

Lema 2

i. *$GL(\mathbb{R}^n, \mathbb{R}^n)$ es un subconjunto abierto de $L(\mathbb{R}^n, \mathbb{R}^n)$.*

ii. *$\mathcal{L}^{-1}$ es una transformación C^∞.*

Demostración

i. La aplicación determinante $\det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n$ (n tiempos) $\to \mathbb{R}$ es una transformación n-lineal (recuérdese que el determinante es lineal en las filas). Así, det se puede ver como un polinomio en n^2 variables, por lo que es continuo; de hecho, es de clase C^∞. Como el conjunto $\{0\}$ es cerrado, $\det^{-1}(\{0\})$ es cerrado. Por lo tanto, $L(\mathbb{R}^n, \mathbb{R}^n) \setminus \det^{-1}(\{0\})$ es abierto. Pero $L(\mathbb{R}^n, \mathbb{R}^n) \setminus \det^{-1}(\{0\})$ es el conjunto de matrices $n \times n$ con determinante distinto de cero, y éstas son precisamente todas las matrices invertibles $GL(\mathbb{R}^n, \mathbb{R}^n)$.

ii. De la expresión explícita para la inversa de una matriz, vemos que $\mathcal{L}^{-1}$ es C^∞. De hecho, la expresión para la inversa de la matriz A es $A^{-1} = (\det A)^{-1}$, donde adj A es una matriz tal que $(\operatorname{adj} A)_{ij} = (-1)^{i+j} \det A(j \mid i)$, donde $\det A(j \mid i)$ denota el determinante de la matriz que se obtiene de A eliminando la j-ésima fila y la i-ésima columna. Como $(\det A)^{-1}$ es una función diferenciable de A, sólo necesitamos mostrar que la aplicación adj : $L(\mathbb{R}^n, \mathbb{R}^n) \to L(\mathbb{R}^n, \mathbb{R}^n)$ que lleva a cada matriz en su adjunta es C^∞. Considerada como una función de $\mathbb{R}^{n^2}$ en $\mathbb{R}^{n^2}$, la adjunta es

simplemente una n^2-upla de funciones como $(\text{adj } A)_{ij} = (-1)^{i+j} \det A(j \mid i)$. Como ya hemos mencionado, una aplicación multilineal de $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_k}$ en $\mathbb{R}^m$ es C^∞. Así, cada una de las n^2 funciones componentes de adj es C^∞. Por lo tanto, la aplicación adj es C^∞. ▼

Demostración del teorema de la función inversa Para mayor claridad, separamos la demostración en varios pasos.

Paso 1 *Simplificación a un caso particular.*

Demostramos el teorema cuando $\mathbf{D}f(x_0)$ es la transformación identidad. Para ver que esto es suficiente para demostrar el caso general, sea $T = \mathbf{D}f(x_0)$. Entonces, T^{-1} existe y es lineal, de modo que $\mathbf{D}(T^{-1})(f(x_0)) = T^{-1}$ y, por la regla de la cadena,

$$\begin{aligned}\mathbf{D}(T^{-1} \circ f)(x_0) &= \mathbf{D}(T^{-1})(f(x_0)) \circ \mathbf{D}f(x_0) = T^{-1} \circ \mathbf{D}f(x_0)\\ &= \text{transformación identidad.}\end{aligned}$$

Si el teorema es cierto para $T^{-1} \circ f$, entonces el teorema también es cierto para f. De hecho, si g es una inversa de $T^{-1} \circ f$, la inversa de f será $g \circ T^{-1}$.

Podemos hacer otra simplificación, a saber, que $x_0 = 0$ y $f(x_0) = 0$. Para ver esto, supóngase que hemos demostrado el teorema para el caso particular $x_0 = 0$ y $f(x_0) = 0$. Para demostrar el caso general a partir de esto, sea $h(x) = f(x + x_0) - f(x_0)$. Entonces $h(0) = 0$ y $\mathbf{D}h(0) = \mathbf{D}f(x_0)$, de modo que $\mathbf{D}h(0)$ es invertible. Si h tiene una inversa cerca de $x = 0$, entonces la inversa de f cerca de x_0 será

$$f^{-1}(y) = h^{-1}(y - f(x_0)) + x_0.$$

En resumen, el paso 1 prueba que *es suficiente demostrar el teorema bajo las hipótesis de que* $x_0 = 0$, $f(x_0) = 0$ *y* $\mathbf{D}f(0)$ *es la identidad.* Supondremos esto en el resto del análisis.

Paso 2 *Aplicación del principio de la aplicación contractiva para obtener una inversa local.*

Con las observaciones anteriores en mente, lo que quisiéramos son dos vecindades de 0 tales que, dado cualquier y de la primera vecindad de 0, exista un único x de la segunda vecindad tal que $f(x) = y$. Para mostrar esto, considérese la función g_y dada por $g_y(x) = y + x - f(x)$. Si para alguna vecindad cerrada de cero ésta es una aplicación contractiva, entonces tiene un único punto fijo, digamos, x, de modo que $x = y + x - f(x)$, o sea x es el único punto perteneciente a la vecindad tal que $f(x) = y$. Para construir esta vecindad, defínase $g(x) = x - f(x)$; entonces $\mathbf{D}g(0) = 0$. Supóngase que g es de clase C^p, $p \geq 1$. Esto significa, en particular, que $\mathbf{D}g$ es una función continua, y por la continui-

dad en 0, existe $r > 0$ tal que $\|x\| < r$ implique $\|\mathbf{D}g_i(x)\| < 1/2n$, donde $g = (g_1, \ldots, g_n)$. Por el teorema del valor medio, dado $x \in \mathbf{D}(0, r)$ existen puntos $c_1, c_2 \ldots, c_n$ en $\mathbf{D}(0, r)$ tales que $g_i(x) = g_i(x) - g_i(0) = \mathbf{D}g_i(c_i)(x - 0) = \mathbf{D}g_i(c_i)(x)$. En consecuencia,

$$\|g(x)\| \leq \sum_{i=1}^{n} |g_i(x)| = \sum_{i=1}^{n} |\mathbf{D}g_i(c_i)(x)| \leq \sum_{i=1}^{n} \|\mathbf{D}g_i(c_i)\|\ \|x\| < \frac{\|x\|}{2} < \frac{r}{2},$$

aplicando la desigualdad de Cauchy-Schwarz a $\mathbf{D}g_i(c_i)(x) = \langle \nabla g_i(c_i), x\rangle$.

Esto muestra que g aplica la bola cerrada de radio r, $B(0, r)$, en la bola cerrada de radio $r/2$, $B(0, r/2)$. Ahora, sea y cualquier elemento de $B(0, r/2)$. La transformación g_y aplica $B(0, r)$ en $B(0, r)$; para ver esto, obsérvese que $\|y\| \leq r/2$ y $x \in B(0, r)$ implican

$$\|g_y(x)\| = \|y + g(x)\| \leq \|y\| + \|g(x)\| < r/2 + r/2 = r.$$

Sean x_1 y x_2 dos puntos cualesquiera en $B(0, r)$. Entonces $\| g_y(x_1) - g_y(x_2) \| = \| g(x_1) - g(x_2) \|$ y, por el teorema del valor medio, como antes, $\| g(x_1) - g(x_2) \| \leq (1/2) \| x_1 - x_2 \|$, de modo que g_y es una aplicación contractiva (con constante $K = 1/2$). El principio de la aplicación contractiva implica entonces que existe un único punto fijo $x \in B(0, r)$ para g_y, y, como ya hemos observado, esto implica que $f(x) = y$. Esto significa que f tiene una inversa $f^{-1} : B(0, r/2) \subset \mathbb{R}^n \to B(0, r) \subset \mathbb{R}^n$.

Paso 3 *La inversa es continua.*

Sean x_1 y $x^2 \in B(0, r)$. Si recordamos la definición de g, obtenemos

$$\|x_1 - x_2\| \leq \|f(x_1) - f(x_2)\| + \|g(x_1) - g(x_2)\| \leq \|f(x_1) - f(x_2)\| + \frac{1}{2}\|x_1 - x_2\|,$$

y por lo tanto, $\| x_1 - x_2 \| \leq 2 \|f(x_1) - f(x_2)\|$. Así, si y_1 e $y_2 \in B(0, 1/2)$, entonces $\|f^{-1}(y_1) - f^{-1}(y_2)\| \leq 2 \| y_1 - y_2 \|$, de modo que f^{-1} es continua.

Paso 4 *Para r suficientemente pequeña, la inversa es diferenciable en* $\mathbf{D}(0, r/2)$.

Aquí tenemos que $\mathbf{D}f(0)$ es invertible y que $\mathbf{D}f : A \subset \mathbb{R}^n \to \mathbb{R}^{n^2}$ es continua, y hemos mostrado que $GL(\mathbb{R}^n, \mathbb{R}^n)$ es abierto en $L(\mathbb{R}^n, \mathbb{R}^n)$. Estos hechos muestran que para todo x en alguna vecindad con centro en 0, $[\mathbf{D}f(x)]^{-1}$ existe. Si esta vecindad no contiene a $\mathbf{D}(0, r/2)$, r se restringe aún más hasta que esto ocurra. Por lo tanto, podemos suponer que $[\mathbf{D}f(x)]^{-1}$ existe para todo $x \in \mathbf{D}(0, r/2)$. Además, podemos suponer que $\| [\mathbf{D}f(x)]^{-1} y \| \leq M \| y \|$ para todo $x \in \mathbf{D}(0, r/2)$ e $y \in \mathbb{R}^n$, por la continuidad de $[\mathbf{D}f(x)]^{-1}$ (véase el ejemplo resuelto **4.4**, capítulo 4).

Para y_1 e $y_2 \in \mathbf{D}(0, r/2)$, $x_1 = f^{-1}(y_1)$ y $x_2 = f^{-1}(y_2)$,

$$\frac{\|f^{-1}(y_1) - f^{-1}(y_2) - [\mathbf{D}f(x_2)]^{-1} \cdot (y_1 - y_2)\|}{\|y_1 - y_2\|}$$

$$= \frac{\|x_1 - x_2 - [\mathbf{D}f(x_2)]^{-1} \cdot (f(x_1) - f(x_2))\|}{\|f(x_1) - f(x_2)\|}$$

$$= \left[\frac{\|x_1 - x_2\|}{\|f(x_1) - f(x_2)\|}\right] \frac{\|\{[\mathbf{D}f(x_2)]^{-1}\}\{\mathbf{D}f(x_2)(x_1 - x_2) - (f(x_1) - f(x_2))\}\|}{\|x_1 - x_2\|}.$$

Usando $\|x_1 - x_2\| \leq 2\,\|f(x_1) - f(x_2)\|$ y $\|\mathbf{D}f(x^2)^{-1}y\| \leq M\,\|y\|$ esta expresión resulta ser

$$\leq 2M \frac{\|\mathbf{D}f(x_2)(x_1 - x_2) - (f(x_1) - f(x_2))\|}{\|x_1 - x_2\|}.$$

La última expresión tiene límite cero cuando $\|x_1 - x_2\| \to 0$, por la diferenciabilidad de f en x_2. Esto muestra que f^{-1} es diferenciable en y_2 con diferencial $[\mathbf{D}f(x_2)]^{-1} = [\mathbf{D}f(f^{-1}(y_2))]^{-1}$.

En el teorema, hacemos $W = \mathbf{D}(0, r/2)$ y $U = f^{-1}(W)$, ambos conjuntos abiertos.

Paso 5 $f^{-1} : \mathbf{D}(0, r/2) \to \mathbb{R}^n$ *es de clase* C^p.

El paso 4 implica que la transformación inversa $f^{-1} : \mathbf{D}(0, r/2) \to \mathbb{R}^n$ es diferenciable en $\mathbf{D}(0, r/2)$ y que $\mathbf{D}f^{-1}(y) = [\mathbf{D}f(f^{-1}(y))]^{-1}$. También hemos mostrado que $f^{-1} : \mathbf{D}(0, r/2) \to \mathbb{R}^n$ es continua: $\mathbf{D}f$ es continua por hipótesis y la transformación de inversión de $GL(\mathbb{R}^n, \mathbb{R}^n)$ (el conjunto de transformaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^n$) a $GL(\mathbb{R}^n, \mathbb{R}^n)$ es continua y, de hecho, C^∞, por el lema **2**. Esto implica que $\mathbf{D}f^{-1}$ es una transformación continua de $\mathbf{D}(0, r/2)$ en $L(\mathbb{R}^n, \mathbb{R}^n)$. Por lo tanto, f^{-1} es de clase C^1.

Analícese de nuevo $\mathbf{D}f^{-1}(y) = [\mathbf{D}f(f^{-1}(y))]^{-1}$ y observemos que como f^{-1} es de clase C^1, $\mathbf{D}f$ es clase C^{p-1} y, como la inversión es C^∞, $\mathbf{D}f^{-1}$ es de clase C^1. Por lo tanto, f^{-1} es de clase C^2. Si continuamos de esta forma por inducción, concluimos finalmente que f^{-1} es de clase C^p. ■

7.2.1 Teorema de la función implícita *Sean* $A \subset \mathbb{R}^n \times \mathbb{R}^m$ *un conjunto abierto y* $F : A \to \mathbb{R}^m$ *una función de clase* C^p *(es decir, supóngase que* F *tiene* p *diferenciales continuas, donde* p *es un entero,* $p \geq 1$*). Supóngase que* $(x_0, y_0) \in A$ *y* $F(x_0, y_0) = 0$. *Fórmese el determinante*

$$\Delta = \begin{vmatrix} \dfrac{\partial F_1}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F_m}{\partial y_1} & \cdots & \dfrac{\partial F_m}{\partial y_m} \end{vmatrix}$$

evaluado en (x_0, y_0), donde $F = (F_1, \ldots, F_m)$. Supóngase que $\Delta \neq 0$. Entonces existen una vecindad abierta $U \subset \mathbb{R}^n$ de x_0, una vecindad V de y_0 en $\mathbb{R}^m$ y una única función $f : U \to V$ tal que

$$F(x, f(x)) = 0$$

para todo $x \in U$. Además, f es de clase C^p.

Demostración Defínase la función $G : A \subset \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^m$ como $G(x, y) = (x, F(x, y))$. Para x cercano a x_0, queremos encontrar $y = f(x)$ tal que $F(x, y) = 0$. La idea es demostrar que G es localmente invertible cerca de (x_0, y_0) y tomar $f(x)$ determinada por $(x, f(x)) = G^{-1}(x, 0)$. Como F es de clase C^p y la transformación identidad es de clase C^∞, se sigue que G es de clase C^p. La matriz de derivadas parciales de G (la matriz jacobiana) es

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & & & \cdots & \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & & \cdots & 1 & 0 & \cdots & 0 \\ \dfrac{\partial F_1}{\partial x_1} & & \cdots & \dfrac{\partial F_1}{\partial x_n} & \dfrac{\partial F_1}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_m} \\ \vdots & & & \vdots & \vdots & & \vdots \\ \dfrac{\partial F_m}{\partial x_1} & & \cdots & \dfrac{\partial F_m}{\partial x_n} & \dfrac{\partial F_m}{\partial y_1} & \cdots & \dfrac{\partial F_m}{\partial y_m} \end{pmatrix}$$

El determinante de esta matriz evaluada en (x_0, y_0) es igual a

$$\Delta = \begin{vmatrix} \dfrac{\partial F_1}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F_m}{\partial y_1} & \cdots & \dfrac{\partial F_m}{\partial y_m} \end{vmatrix}$$

evaluado en (x_0, y_0). Por hipótesis, $JG(x_0, y_0) \neq 0$ y entonces, por el teorema de la función inversa, existe un conjunto abierto W que contiene a $(x_0, 0)$ y un conjunto abierto S que contiene a (x_0, y_0) tales que $G(S) = W$ y G tiene una inversa C^p, $G^{-1} : W \to S$. Por definición de conjunto abierto, existen conjuntos abiertos $U \subset \mathbb{R}^n$ y $V \subset \mathbb{R}^m$, con $x_0 \in U$ e $y_0 \in V$ tales que $U \times V \subset S$. Sea $G(U \times V) = Y \subset W$. Así, G: $U \times V \to Y$ es un difeomorfismo C^p (esto significa que G es de clase C^p y que tiene inversa $G^{-1} : Y \to U \times V$, también de clase C^p). Ahora, G^{-1} es de la forma $G^{-1}(x, w) : (x, H(x, w))$, donde H es

una función C^p de Y en V, pues G es de esta forma, como es fácil de ver. Sea $\pi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ dada por $\pi(x, y) = y$, de modo que $F(x, H(x, w)) = \pi \circ G(x, H(x, w)) = \pi \circ G \circ G^{-1}(x, w) = w$. Como G^{-1} es de la forma $G^{-1}(x, w) = (x, H(x, w))$, $x \in U$ si $(x, w) \in Y$. Defínase $f : U \to V$ por $f(x) = H(x, 0)$. Como $F(x, H(x, w)) = w$, obtenemos $F(x, f(x)) = 0$, y como H es de clase C^p, f también es de clase C^p. Por el teorema de la función inversa, $H(x, w)$ está determinada en forma única. Como f debe estar dada por $H(x, 0)$, f también es única. ■

7.3.1 Teorema de rectificación del dominio *Sean $A \subset \mathbb{R}^n$ un conjunto abierto y $f : A \to \mathbb{R}$ una función de clase $C^p, p \geq 1$. Sea $x_0 \in A$ y supóngase que $f(x_0) = 0$ y que* $\mathbf{D}f(x_0) \neq 0$. *Entonces existen un conjunto abierto U, un conjunto abierto V que contiene a x_0 y una función $h : U \to V$ de clase C^p, con inversa $h^{-1} : V \to U$ de clase C^p, tales que*

$$f(h(x_1, \ldots, x_n)) = x_n.$$

Demostración Como $\mathbf{D}f(x_0) \neq 0$, existe i tal que $(\partial f/\partial x_i)(x_0) \neq 0$. Defínase $g : \mathbb{R}^n \to \mathbb{R}^n$ como $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_{i-1}, x_n, x_{i+1}, \ldots, x_{n-1}, x_i)$. La aplicación de permutación g es lineal y por lo tanto es C^∞; como f es de clase C^p, la regla de la cadena muestra que $f \circ g$ es de clase C^p. Así, $(\partial(f \circ g)/\partial x_n)(g^{-1})(x_0) = \partial f(x_0)/\partial x_i \neq 0$, lo que implica que $f \circ g$ es una función del tipo descrito en las hipótesis del teorema de la función implícita, con $m = 1$. Como en la demostración de **7.2.1**, si definimos $G : A \subset \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}^{n-1} \times \mathbb{R}$ como $G(x, y) = (x, f \circ g(x, y))$, existen conjuntos abiertos $W \subset \mathbb{R}^n$ y $U \subset \mathbb{R}^n$ con $x_0 \in W$ y $(x_0^1, \ldots, x_0^{n-1}, 0) \in U$ (donde $x_0 = (x_0^1, \ldots, x_0^n)$) tales que $G : W \to U$ tiene una inversa $G^{-1} : U \to W$ de clase C^p. Ahora, $(f \circ g) \circ G^{-1}(x_1, \ldots, x_n) = (\pi \circ G) \circ G^{-1}(x_1, \ldots, x_n) = (\pi \circ G) \circ G^{-1}(x) = (x_1, \ldots, x_n) = x_n$, donde $\pi : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ es la proyección en la última coordenada. Defínase $V = g(W)$ y $h : U \to V$ como $h = g \circ G^{-1}$. Entonces h es una función C^p con inversa C^p, pues g y G^{-1} tienen esta propiedad y $f(h(x_1, \ldots, x_n)) = x_n$. ■

Es posible demostrar un teorema más general que **7.3.1**, usando una técnica similar. Es decir, si $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, $n \geq m$ y $\mathbf{D}f(x_0)$ como transformación lineal tiene rango m, entonces es posible hacer que f se parezca a una proyección en los últimos m factores al componerla con una función suave de inversa suave. En el ejercicio 18 del final del capítulo estableceremos esto de forma precisa y daremos una sugerencia para la demostración. Obsérvese aquí que el rango tiene dimensión menor o igual que la del dominio. En el siguiente teorema ocurre lo contrario.

7.4.1 Teorema de rectificación de la imagen *Sean $A \subset \mathbb{R}^p$ un conjunto abierto y $f : A \to \mathbb{R}^n$ una función de clase C^r, $r \geq 1$, y supóngase que $p \leq n$. Sea $x_0 \in A$ y*

supóngase que el rango de $\mathbf{D}f(x_0)$ *es p. Entonces existen conjuntos abiertos U y V en* $\mathbb{R}^n$ *tales que* $f(x_0) \in U$ *y existe una función* $g : U \to V$ *de clase* C^r *con inversa* $g^{-1} : V \to U$ *también de clase* C^r*, tales que* $g \circ f(x_1, \ldots, x_p) = (x_1, \ldots, x_p, 0, \ldots, 0)$ *para todo* $(x_1, \ldots, x_p) \in V$.

Demostración Como $\mathbf{D}f(x_0)$ tiene rango p, alguna submatriz $p \times p$ de $\mathbf{D}f(x_0)$ tiene determinante distinto de cero. Podemos renombrar las variables, en caso necesario, para suponer que

$$\begin{vmatrix} \dfrac{\partial f^1}{\partial x_1} & \cdots & \dfrac{\partial f^1}{\partial x_p} \\ \vdots & & \vdots \\ \dfrac{\partial f^p}{\partial x_1} & \cdots & \dfrac{\partial f^p}{\partial x_p} \end{vmatrix} \neq 0,$$

donde $f = (f^1, \ldots, f^n)$. Defínase $\varphi : A \times \mathbb{R}^{n-p} \to \mathbb{R}^n$ como $\varphi(x, y) = f(x) + (0, y)$. Entonces la matriz de $\mathbf{D}\varphi$ es

$$\begin{pmatrix} \dfrac{\partial f^1}{\partial x_1} & \cdots & \dfrac{\partial f^1}{\partial x_p} & 0 & \cdots & 0 \\ \vdots & & \vdots & & & \\ \dfrac{\partial f^p}{\partial x_1} & \cdots & \dfrac{\partial f^p}{\partial x_p} & 0 & \cdots & 0 \\ \dfrac{\partial f^{p+1}}{\partial x_1} & \cdots & \dfrac{\partial f^{p+1}}{\partial x_p} & 1 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f^n}{\partial x_1} & \cdots & \dfrac{\partial f^n}{\partial x_p} & 0 & \cdots & 1 \end{pmatrix}$$

de modo que

$$J\varphi(x_0, 0) = \begin{vmatrix} \dfrac{\partial f^1}{\partial x_1} & \cdots & \dfrac{\partial f^1}{\partial x_p} \\ \vdots & & \vdots \\ \dfrac{\partial f^p}{\partial x_1} & \cdots & \dfrac{\partial f^p}{\partial x_p} \end{vmatrix} \neq 0.$$

Por el teorema de la función inversa, existen un conjunto abierto U en torno a $f(x_0)$, un conjunto abierto V en torno a $(x_0, 0)$ y una función $g : U \to V$ de clase C^r tales que $g = \phi^{-1}$. Entonces $g(f(x)) = g(f(x)) + (0, 0) = (x, 0)$, como se pedía. ■

7.4.2 Teorema del rango *Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^N$ (donde A es abierto en $\mathbb{R}^n$) una función C^r tal que $\mathbf{D}f(x)$ tiene rango m para todo x en una vecindad de $x_0 \in A$. Entonces existen un conjunto abierto $U \subset \mathbb{R}^n$, un conjunto abierto $V \subset \mathbb{R}^n$ con $x_0 \in V$ y una función $h : U \to V$ de clase C^r con inversa $h^{-1} : V \to U$ de clase C^r tal que $f \circ h$ depende solamente de $x_1, \ldots, x_m$. Es decir, $(f \circ h)(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n) = \tilde{f}\,(x_1, \ldots, x_m)$ para alguna función C^r, $\tilde{f}$.*

Demostración Sea N_0 el núcleo de $\mathbf{D}f(x_0)$; es decir, sea $N_0 = \{y \in \mathbb{R}^n \mid \mathbf{D}f(x_0) \cdot y = 0\}$ (un subespacio de $\mathbb{R}^n$ de dimensión $n - m$) y sea M un complemento m-dimensional de N_0 en $\mathbb{R}^n$, es decir, sea $M \cap N_0 = \{0\}$ y $\{x + y \mid x \in M, y \in N_0\} = \mathbb{R}^n$. Sea $c_1, \ldots, c_m$ una base de M y $c_{m+1}, \ldots, c_n$ una base de N_0. Cada $x \in \mathbb{R}^n$ se puede escribir de forma única como $x = \psi_1(x)c_1 + \cdots + \psi_n(x)c_n$. Defínase $G(x) = (0, \ldots, 0, \psi_{m+1}(x), \ldots, \psi_n(x))$. Entonces G es lineal y, por lo tanto, suave. Como $\mathbf{D}f(x_0)$ tiene rango m, $\mathbf{D}f(x_0)(\mathbb{R}^n)$ es un subespacio m-dimensional P de $\mathbb{R}^N$. Además, el conjunto $\{d_i = \mathbf{D}f(x_0)c_i \mid 1 \le i \le m\}$ es una base de P. Cualquier $x \in \mathbb{R}^N$ se puede escribir en forma única como

$$x = \varphi_1(x)d_1 + \cdots + \varphi_m(x)d_m + \varphi_{m+1}(x)d_{m+1} + \cdots + \varphi_N(x)d_N,$$

donde $d_1, \ldots, d_N$ es una base de $\mathbb{R}^N$, con $d_1, \ldots, d_m$ la base de P que acabamos de dar. Defínase $H : \mathbb{R}^N \to \mathbb{R}^n$ como $H(x) = (\varphi_1(x), \ldots, \varphi_m(x), 0, \ldots, 0)$.

Sea $g(x) = H(f(x) + G(x))$. Entonces g aplica $\mathbb{R}^n$ en $\mathbb{R}^n$, y como H y G son lineales, tenemos que $\mathbf{D}g(x_0) \cdot s = \mathbf{D}H(f(x_0)) \circ \mathbf{D}f(x_0)(s) + \mathbf{D}G(x_0)(s) = H(\mathbf{D}f(x_0)(s)) + G(s)$. Si escribimos la matriz de la transformación lineal $\mathbf{D}g(x_0)$ en términos de las bases $c_1, \ldots, c_n$ y la base canónica, obtenemos la matriz identidad. Por lo tanto, $\mathbf{D}g(x_0)$ es invertible. Podemos usar el teorema de la función inversa para encontrar un conjunto abierto U en torno a $H(f(x_0)) + G(x_0)$, un conjunto abierto V en torno a x_0 y una función inversa suave $g^{-1} : U \to V$. Para cada $x \in V$, $\mathbf{D}g(x)$ es invertible; es decir, $\mathbf{D}g(x)$ es una transformación lineal inyectiva de $\mathbb{R}^n$ sobre $\mathbb{R}^n$. Podemos suponer que rango$\{\mathbf{D}f(x)\} = m$ para todo $x \in A$ (en caso contrario, restringimos f a una vecindad todavía más pequeña de x_0). Para $x \in A$, $\mathbf{D}f(x)(\mathbb{R}^n)$ es un subespacio m-dimensional P_x de $\mathbb{R}^n$. Para $s \in M$, $\mathbf{D}g(x) \cdot s = H(\mathbf{D}f(x) \cdot s) + G(s) = H(\mathbf{D}f(x) \cdot s)$. Así, si $x \in V$, $\mathbf{D}f(x)$ restringida a M es una transformación lineal inyectiva de M sobre P_x. La suprayectividad de la transformación proviene del hecho de que M y P_x tienen ambos dimensión m. Análogamente, H debe ser una transformación lineal inyectiva de P_x sobre $\mathbb{R}^m$. Denotamos la inversa de esta aplicación como $L_x : \mathbb{R}^m \to P_x$.

Sea $h = g^{-1} : U \to V$; mostraremos que $f \circ h(x_1, \ldots, x_n)$ no depende de $x_{m+1}, \ldots, x_n$. Para esto, podemos suponer que U es una bola. Basta mostrar que $\mathbf{D}_2 f_1 = 0$, donde $\mathbf{D}_2 f_1$ es la derivada de $f_1 = f \circ h$ restringida a $\{0\} \times \mathbb{R}^{n-m}$; es decir, mostraremos que $\partial f_1/\partial x_i = 0$, $i = m + 1, \ldots, n$. Se sigue que f_1 es constante con respecto a $x_{m+1}, \ldots,$

x_n. Ahora, $f = f_1 \circ g$, de modo que

$$\begin{aligned}\mathbf{D}f(x)\cdot y &= \mathbf{D}f_1(g(x))\cdot \mathbf{D}g(x)\cdot y\\ &= \mathbf{D}_1 f_1(g(x))\cdot H(\mathbf{D}f(x)\cdot y) + \mathbf{D}_2 f_1(g(x))\cdot G(y).\end{aligned} \tag{1}$$

Como G es una transformación de $\mathbb{R}^n$ sobre $\{0\}\times\mathbb{R}^{n-m}$, basta mostrar que $\mathbf{D}_2 f_1(g(x)) \circ G(x) = 0$ para $y \in \mathbb{R}^n$. Regresamos a la ecuación (1) y usamos que $L_x \circ H$ = identidad para obtener

$$\begin{aligned}\mathbf{D}_2 f_1(g(x)) \circ G(y) &= L_x \circ H(\mathbf{D} f(x)\cdot y) - \mathbf{D}_1 f_1(g(x)) \circ H(\mathbf{D}f(x)\cdot y)\\ &= (L_x - \mathbf{D}_1 f_1(g(x))) \circ H(\mathbf{D}f(x)\cdot y)\end{aligned} \tag{2}$$

para todo $y \in \mathbb{R}^n$. Ahora, $L_x - \mathbf{D}_1 f_1(g(x))$ está definida en $\mathbb{R}^m \times \{0\}$ y $H \circ \mathbf{D}f(x)$ aplica M sobre $\mathbb{R}^m \times \{0\}$. Por lo tanto, para mostrar que $L_x - \mathbf{D}_1 f_1(g(x)) = 0$, basta mostrar que

$$(L_x - \mathbf{D}_1 f_1(g(x))) \circ H(\mathbf{D}f(x)\cdot y) = 0$$

para $y \in M$. Esto se debe a que para $y \in M$, $G(y) = 0$, de modo que $\mathbf{D}_2 f_1(g(x)) \circ G(y) = 0$. Por lo tanto, $L_x - \mathbf{D}_1 f_1(g(x))$ es idénticamente nula, de donde $\mathbf{D}_2 f_1(g(x)) = 0$. ■

7.4.3 Corolario *Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^N$ (donde A es abierto en $\mathbb{R}^n$) una función de clase C^r, $r \geq 1$, tal que $\mathbf{D}f(x)$ tenga rango m para todo x en una vecindad de $x_0 \in A$. Entonces existen un conjunto abierto $U_1 \subset \mathbb{R}^n$, un conjunto abierto $U_2 \subset \mathbb{R}^n$ tal que $x_0 \in U_2$, un conjunto abierto V_1 en torno a $f(x_0)$, un conjunto abierto $V_2 \subset \mathbb{R}^N$, y funciones $h : U_1 \to U_2$ y $g : V_1 \to V_2$ de clase C^r con inversas de clase C^r tales que $(g \circ f \circ h)(x_1, \ldots, x_n) = (x_1, \ldots, x_m, 0, \ldots, 0)$.*

Demostración Por el teorema **7.4.2**, existe una función C^r, $h : U \to V$, con inversa C^r, tal que

$$(f \circ h)(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n) = \tilde{f}(x_1, \ldots, x_m)$$

para alguna $\tilde{f} : W \subset \mathbb{R}^m \to \mathbb{R}^N$. Como h es invertible, $\mathbf{D}\tilde{f}$ tiene rango m, y por el teorema de rectificación de la imagen, existe una función C^r invertible g_0 tal que

$$(g_0 \circ \tilde{f})(x_1, \ldots, x_m) = (x_1, \ldots, x_m, 0, \ldots, 0).$$

Defínase g en $\mathbb{R}^n$ como $g(x_1, \ldots, x_n) = (g_0(x_1, \ldots, x_m), x_{m+1}, \ldots, x_n)$. Entonces g también es C^r e invertible, y tenemos

$$(g \circ f) \circ h(x_1, \ldots, x_n) = (x_1, \ldots, x_m, 0, \ldots, 0). \quad ■$$

7.5.3 Teorema *Sea $f : [-a, a] \times B(x_0, r) \to \mathbb{R}^n$ una aplicación continua. Sea* $C = \sup\{\|f(t, x)\| \mid (t, x) \in [-a, a] \times B(x_0, r)\}$. *Supóngase que existe K constante tal que*

$$\|f(t,x) - f(t,y)\| \leq K\|x - y\| \qquad \text{(L)}$$

para todo $t \in [-a, a]$, $x, y \in B(x_0, r)$. Sea $b <$ mín$\{a, r/C, 1/K\}$. Entonces existe una única transformación continuamente diferenciable $x : [-b, b] \to B(x_0, r) \subset \mathbb{R}^n$ tal que

$$x(0) = x_0 \quad (\textit{condición inicial}) \quad y \quad \frac{dx}{dt} = f(t, x(t)).$$

Demostración La ecuación diferencial y la condición inicial $x(0) = x_0$ son equivalentes, por el teorema fundamental del cálculo, a la condición

$$x(t) = x_0 + \int_0^t f(s, x(s))\, ds$$

para una función continua $x(t)$. Considérese $C([-b, b], \mathbb{R}^n)$, del que sabemos (por §**5.4**) que es un espacio métrico completo. Sea

$$A = \{\varphi \in C([-b, b], \mathbb{R}^n) \mid \varphi(0) = x_0 \quad y \quad \varphi(t) \in B(x_0, r)\}.$$

Entonces $A \subset C([-b, b], \mathbb{R}^n)$ es cerrado (¿por qué?) y, por ello, A es también un espacio métrico completo. Aplicaremos el principio de la aplicación contractiva a este espacio A. Defínase $F : A \to A$ como

$$F(\varphi)(t) = x_0 + \int_0^t f(s, \varphi(s))\, ds.$$

Aquí, la expresión $\int_0^t f(s, \varphi(s))ds$ se obtiene al integrar cada componente de f; el resultado es un vector. La desigualdad general

$$\left\|\int_0^t f(s, \varphi(s))\, ds\right\| \leq \int_0^t \|f(s, \varphi(s))\|\, ds$$

es análoga al resultado similar para el caso de funciones reales, y lo aceptaremos como un hecho probado (véase §**4.8**). En primer lugar, debemos mostrar que $F(\varphi) \in A$. La función $F(\varphi)$ es continua, pues

$$\|F(\varphi)(\tau) - F(\varphi)(t)\| = \left\|\int_\tau^t f(s, \varphi(s))\, ds\right\| \leq \int_t^\tau \|f(s, \varphi(s))\|\, ds \leq C|\tau - t|$$

para $-b \leq t < \tau \leq b$. Así, $F(\varphi) \in C([-b, b], \mathbb{R}^n)$. Además, $F(\varphi)(0) = x_0$, y para todo $t \in [-b, b]$,

$$\|F(\varphi)(t) - x_0\| = \left\|\int_0^t f(s, \varphi(s))\, ds\right\| \leq \operatorname{sign}(t) \int_0^t \|f(s, \varphi(s))\, ds\| \leq b \cdot C < r,$$

ya que $b < r/C$. El factor sign(t) tiene en cuenta la posibilidad de que $t < 0$. Así, $F(\varphi)(t) \in B(x_0, r)$, de modo que $F(\varphi) \in A$. A continuación, para $\varphi, \psi \in A$,

$$\begin{aligned}\|F(\varphi) - F(\psi)\| &= \sup_{-b\le t\le b} \|F(\varphi)(t) - F(\psi)(t)\| \\ &= \sup_{-b\le t\le b} \left\| \int_0^t f(s, \varphi(s)) - f(s, \psi(s))\, ds \right\| \\ &\le \sup_{-b\le t\le b} \operatorname{sign}(t) \int_0^t \|f(s, \varphi(s)) - f(s, \psi(s))\|\, ds \\ &\le \sup_{-b\le t\le b} \operatorname{sign}(t) \int_0^t K\|\varphi(s) - \psi(s)\|\, ds \\ &\le \sup_{-b\le t\le b} K \operatorname{sign}(t) \int_0^t \|\varphi - \psi\|\, ds \le Kb\|\varphi - \psi\|\end{aligned}$$

donde $Kb < 1$. En consecuencia, si hacemos $k = b \cdot K < 1$, entonces $d(F(\varphi), F(\psi)) \le kd(\varphi, \psi)$, por lo que F es una contracción y por tanto tiene un único punto fijo: $x = F(x)$. Este punto fijo $x(t)$ es la solución única que buscábamos. ■

El esquema de iteración mencionado en el texto viene al caso pues, como vimos en la demostración del teorema de la aplicación contractiva, el único punto fijo es el límite $F^n(\varphi)$ cuando $n \to \infty$ para cualquier $\varphi \in A$, tal que $\varphi(t) \equiv x_0$.

7.6.1 Lema de Morse *Sean $A \subset \mathbb{R}^n$ un conjunto abierto y $f : A \to \mathbb{R}$ una función suave (infinitamente derivable). Supóngase que $\mathbf{D}f(x_0) = 0$ y que la hessiana de f en x_0 no es singular. Entonces existen una vecindad U de x_0 y una vecindad V de 0 en $\mathbb{R}^n$, así como una aplicación suave $g : V \to U$ con inversa suave tales que $f \circ g = h$ tiene la forma*

$$h(y) = f(x_0) - [y_1^2 + y_2^2 + \cdots + y_\lambda^2] + [y_{\lambda+1}^2 + \cdots + y_n^2],$$

donde λ es algún entero fijo entre 0 y n.

Demostración No perdemos generalidad si suponemos que $x_0 = 0$ y que $f(x_0) = 0$. Escribimos

$$f(x_1, \dots, x_n) = \int_0^1 \frac{df(tx_1, \dots, tx_n)}{dt}\, dt = \int_0^1 \sum_{i=1}^n x_i \frac{\partial f}{\partial x_i}(tx_1, \dots, tx_n)\, dt.$$

Si hacemos

$$g_i(x_1, \dots, x_n) = \int_0^1 \frac{\partial f}{\partial x_i}(tx_1, \dots, tx_n)\, dt,$$

entonces

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i g_i(x_1, \ldots, x_n).$$

Como $x_0 = 0$ es un punto crítico, $(\partial f/\partial x_i)(0) = g_i(0) = 0$. Además, las funciones g_i son suaves, así que sólo hay que justificar la derivación bajo el signo de integral (se puede aceptar esta afirmación por el momento, o consultar más adelante el ejemplo resuelto **9.2** al final del capítulo 9 para una justificación detallada).

Como $g_i(0) = 0$, podemos aplicar el mismo procedimiento nuevamente para escribir

$$g_i(x_1, \ldots, x_n) = \sum_{j=1}^{n} x_j h_{ij}(x_1, \ldots, x_n)$$

para ciertas funciones suaves h_{ij} y, en consecuencia,

$$f(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} x_i x_j h_{ij}(x_1, \ldots, x_n).$$

Podemos suponer que $h_{ij} = h_{ji}$, si reemplazamos h_{ij} por $\tilde{h}_{ij} = (1/2)(h_{ij} + h_{ji})$, en caso necesario; esto no altera la expresión de f. Obsérvese que en cero, $\partial^2 f/\partial x_i \partial x_j = 2h_{ij}(0)$, de modo que $h_{ij}(0)$ no es singular.

Ahora, f se escribe de manera análoga a una forma cuadrática. Lo que queremos hacer es "diagonalizarla". Procedemos por inducción. Supóngase que existen coordenadas $u_1, \ldots, u_n$ en una vecindad U_1 de 0 tales que

$$f = \pm(u_1)^2 \pm \cdots \pm (u_{r-1})^2 + \sum_{i,j \geq r} u_i u_j H_{ij}(u_1, \ldots, u_n) \tag{1}$$

sobre U_1, donde $r \geq 1$ y las funciones H_{ij} son simétricas. Ya hemos establecido esto para $r = 1$ en el párrafo anterior (la frase "coordenadas $u_1, \ldots, u_n$" significa, como en el texto, que $(u_1, \ldots, u_n)$ son funciones invertibles de $(x_1, \ldots, x_n)$).

Podemos hacer un cambio lineal de coordenadas en $u_r, \ldots, u_n$ para diagonalizar

$$\sum_{i,j \geq r} u_i u_j H_{ij}(0).$$

En particular, como $H_{ij}(0)$ no es singular, los términos de la diagonal son distintos de cero. Así, podemos suponer que $H_{rr}(0) \neq 0$. Sea $g(u_1, \ldots, u_n) = |H_{rr}(u_1, \ldots, u_n)|^{1/2}$; en alguna vecindad más pequeña $U_2 \subset U_1$ de 0, g será una función C^∞ distinta de cero. Defínase

$$\begin{cases} V_i = u_i, \quad i \neq r \\ V_r(u_1, \ldots, u_n) = g(u_1, \ldots, u_n) \left[u_r + \sum_{i>r} \dfrac{u_i H_{ir}(u_1, \ldots, u_n)}{H_{rr}(u_1, \ldots, u_n)} \right] \end{cases} \tag{2}$$

La matriz jacobiana en 0 es

$$\frac{\partial(V_1,\ldots,V_n)}{\partial(u_1,\ldots,u_n)} = \begin{pmatrix} 1 & 0 & & \cdots & & & 0 \\ 0 & 1 & & & & & \\ \vdots & & \ddots & & & & \vdots \\ & & & 1 & & & 0 \\ \dfrac{\partial V_r}{\partial x_1} & & \cdots & g(0) & \cdots & & \dfrac{\partial V_r}{\partial x_n} \\ & & & & 1 & & \\ \vdots & & & & & \ddots & \vdots \\ 0 & & & & & & 0 \\ 0 & & & \cdots & & 0 & 1 \end{pmatrix},$$

que no es singular. En consecuencia, por el teorema de la función inversa, $(u_1, \ldots, u_n) \mapsto (V_1, \ldots, V_n)$ es una aplicación C^∞ con inversa C^∞ en alguna vecindad más pequeña U_3 de 0. En otras palabras, $(V_1, \ldots, V_n)$ servirán como coordenadas. Considérese ahora

$$u_r u_r H_{rr}(u_1,\ldots,u_n) + 2\sum_{j=r+1}^{n} u_j u_r H_{jr}(u_1,\ldots,u_n), \tag{3}$$

que consta de los términos de la ecuación (1) con i o $j = r$. Aquí hemos usado la simetría de H_{ij}. Al comparar la ecuación (1) con la ecuación (2), vemos que la expresión (3) es igual a

$$\pm V_r V_r - \frac{1}{H_{rr}}\left[\sum_{i>r} u_r H_{ir}(u_1,\ldots,u_n)\right]^2,$$

donde el signo más o menos proviene del hecho de que $H_{rr} = \pm g^2$, donde usamos + si H_{rr} es positivo y − si H_{rr} es negativo. De esto vemos que la ecuación (1) se escribe

$$f = \sum_{i\le r} \pm(V_i)^2 + \sum_{i,j>r} V_i V_j \tilde{H}_{ij}(V_1,\ldots,V_n)$$

para unas nuevas $\tilde{H}_{ij}$ simétricas. Así, hemos pasado inductivamente de r a $r + 1$ en la ecuación (1). Por lo tanto, es cierta para $r = n + 1$, lo que demuestra el teorema. ■

7.7.1 Teorema del multiplicador de Lagrange *Sean $f : U \subset \mathbb{R}^n \to \mathbb{R}$ y $g : U \subset \mathbb{R}^n \to \mathbb{R}$ funciones C^1 dadas. Sean $x_0 \in U$, $g(x_0) = c_0$ y $S = g^{-1}(c_0)$ el conjunto de nivel de g con valor c_0. Supóngase que $\nabla g(x_0) \neq 0$. Si $f \mid S$, lo que denota la restricción*

de f a S (es decir, a aquellos $x \in U$ que satisfagan $g(x) = c_0$), tiene un máximo o un mínimo en x_0, entonces existe un número real λ tal que

$$\nabla f(x_0) = \lambda \nabla g(x_0).$$

Demostración Lo único que falta en el bosquejo de demostración dado en **§7.7** es que necesitamos saber que si $v \perp \nabla g(x_0)$, entonces $v = c'(0)$ para una curva C^1, $c(t)$, en S, con $c(0) = x_0$.

Podemos establecer esto de la forma siguiente. Por el teorema de rectificación del dominio, existe un cambio de coordenadas h tal que $g(h(x_1, \ldots, x_n))$. Así, $h^{-1}(S)$ es el plano coordenado $x_n = c_0$. Sea $w = \mathbf{D}h^{-1}(x_0) \cdot v$. Afirmamos que la última coordenada de w es cero; es decir, que w está en el plano $x_n = c_0$. Para mostrar esto, sea $e_n = (0, 0, \ldots, 1)$; mostraremos que $\langle w, e_n \rangle = 0$. Por la regla de la cadena, $g(h(x_1, \ldots, x_n)) = x_n$ implica

$$\langle \nabla g(x_0), \mathbf{D}h(y_0) \cdot w \rangle = \langle w, e_n \rangle,$$

donde $h(y_0) = x_0$. Pero el miembro izquierdo es $\langle \nabla g(x_0), v \rangle = 0$. Sea $c(t) = h(y_0 + tw)$, que está en S, y $c(0) = x_0$; por la regla de la cadena, $c'(0) = v$. La demostración se puede terminar ahora como en el texto. ■

He aquí otra demostración. Supóngase que $(\partial g/\partial x_n)(x_0) \neq 0$. Entonces existe una función h tal que $g(x_1, \ldots, x_{n-1}, h(x_1, \ldots, x_{n-1})) = 0$ para todo $(x_1, \ldots, x_{n-1})$ en un conjunto abierto $U \subset \mathbb{R}^{n-1}$. Sea $k(x_1, \ldots, x_{n-1}) = f(x_1, \ldots, x_{n-1}, h(x_1, \ldots, x_{n-1}))$. Queremos maximizar k. Si k tiene un máximo o un mínimo en un punto, entonces

$$\frac{\partial k}{\partial x_i} = \frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial x_n}\frac{\partial h}{\partial x_i} = 0.$$

Pero

$$\frac{\partial g}{\partial x_i} + \frac{\partial g}{\partial x_n}\frac{\partial h}{\partial x_i} = 0,$$

de modo que

$$\frac{\partial f}{\partial x_i} = \frac{\dfrac{\partial f}{\partial x_n}}{\dfrac{\partial g}{\partial x_n}}\frac{\partial g}{\partial x_i}.$$

Tomando

$$\lambda = \frac{\dfrac{\partial f}{\partial x_n}}{\dfrac{\partial g}{\partial x_n}}$$

obtenemos el resultado. ■

Ejemplos resueltos del capítulo 7

Ejemplo 7.1 (Regla del producto para jacobianos) *Sean $f : A \subset \mathbb{R}^n \to \mathbb{R}^n$, $g : B \subset \mathbb{R}^n \to \mathbb{R}^n$ y $f(A) \subset B$. Muéstrese que para $x \in A$,*

$$J_{g \circ f}(x) = J_g(f(x)) \cdot J_f(x)$$

(producto de números reales).

Solución Por la regla de la cadena,

$$\mathbf{D}(g \circ f)(x) = \mathbf{D}g(f(x)) \circ \mathbf{D}f(x),$$

lo que se puede interpretar como una composición de transformaciones lineales o como un producto de matrices. Como el determinante de un producto de matrices es el producto de los determinantes, obtenemos el resultado pedido. ♦

Ejemplo 7.2 *Considérense las ecuaciones $u = f_1(x, y)$ y $v = f_2(x, y)$. Muéstrese que son invertibles cerca de (x_0, y_0) sii*

$$\Delta = \frac{\partial f_1}{\partial x}\frac{\partial f_2}{\partial y} - \frac{\partial f_1}{\partial y}\frac{\partial f_2}{\partial x}$$

no se anula en (x_0, y_0). Si $x(u, v)$, $y(u, v)$ son las soluciones, muéstrese que

$$\frac{\partial x}{\partial u} = \frac{1}{\Delta}\frac{\partial v}{\partial y}, \qquad \frac{\partial x}{\partial v} = -\frac{1}{\Delta}\frac{\partial u}{\partial y},$$

$$\frac{\partial y}{\partial u} = -\frac{1}{\Delta}\frac{\partial v}{\partial x}, \qquad \frac{\partial y}{\partial v} = \frac{1}{\Delta}\frac{\partial u}{\partial x}.$$

Solución Éste es un caso particular del teorema de la función inversa, para $n = 2$, donde Δ es el determinante jacobiano. La matriz de derivadas de las soluciones es la inversa de la matriz de derivadas de f_1, f_2. Como la inversa de la matriz

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

es

$$\frac{1}{\Delta}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

donde $\Delta = ad - bc$, obtenemos el resultado. ♦

Ejemplo 7.3 *Sean $A \subset \mathbb{R}^n$ un conjunto abierto y $f : A \subset \mathbb{R}^n \to \mathbb{R}^n$ una función inyectiva continuamente diferenciable tal que $Jf(x) = \det(\mathbf{D}f(x)) \neq 0$ para todo $x \in A$. Muéstrese que $f(A)$ es un conjunto abierto y que $f^{-1} : f(A) \to A$ es diferenciable.*

Solución Sea $y \in f(A)$ y supóngase que $y = f(x)$. Como f es continuamente diferenciable y $\mathbf{D}f(x)$ tiene determinante distinto de cero, el teorema de la función inversa nos dice que existen vecindades abiertas U de x y V de y tales que $f|U$ (la restricción de f a U) es un difeomorfismo C^1 (es decir, tiene una inversa C^1) de U sobre V. Por lo tanto, $V \subset f(A)$, de modo que $f(A)$ es abierto. Ahora bien, $(f|U)^{-1} = f^{-1}|f(U)$ y $(f|U)^{-1}$ es diferenciable en y, de modo que f^{-1} es diferenciable en y. Por lo tanto, f^{-1} es diferenciable en $f(A)$. ♦

Ejemplo 7.4 *Considérense las siguientes ecuaciones:*

$$\begin{cases} x^2 - yu = 0, \\ xy + uv = 0. \end{cases}$$

Usando el teorema de la función implícita, descríbanse las condiciones en las cuales se pueden despejar u y v en estas ecuaciones. Después, resuélvanse las ecuaciones directamente y verifíquense dichas condiciones.

Solución Defínase $f_1 : \mathbb{R}^4 \to \mathbb{R}$ como $f_1(x, y, u, v) = x^2 - yu$, y $f_2 : \mathbb{R}^4 \to \mathbb{R}$ como $f_2(x, y, u, v) = xy + uv$. Sea $f : \mathbb{R}^4 \to \mathbb{R}^2$ dada por $f = (f_1, f_2)$: entonces, f es una función suave. La matriz

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial u} & \dfrac{\partial f_1}{\partial v} \\ \dfrac{\partial f_2}{\partial u} & \dfrac{\partial f_2}{\partial v} \end{pmatrix} = \begin{pmatrix} -y & 0 \\ v & u \end{pmatrix}$$

tiene determinante $-yu$. Si (x_0, y_0, u_0, v_0) es tal que $y_0u_0 \neq 0$, entonces se satisfacen las hipótesis del teorema de la función implícita, por lo que existen vecindades A de (x_0, y_0) y B de (u_0, v_0) y una única función continuamente diferenciable $g : A \to B$ tal que $f(x, y, g(x, y)) = 0$ para todo $(x, y) \in A$. Si hacemos $u = g_1$ y $v = g_2$ (donde $g = (g_1, g_2)$), entonces u y v son las soluciones de las ecuaciones simultáneas. Estas ecuaciones tienen soluciones únicas u y v en las vecindades en torno a (x_0, y_0) y (u_0, v_0) que satisfacen las ecuaciones siempre que $y_0u_0 \neq 0$, lo que es equivalente a pedir que $x_0 \neq 0$ e $y_0 \neq 0$, pues $f(x_0, y_0, u_0, v_0) = 0$, o que $x_0^2 - y_0u_0 = 0$ y $x_0y_0 + u_0v_0 = 0$.

Mediante un cálculo directo, las soluciones son $u = x^2/y$ y $v = -y^2/x$, que son válidas excepto cerca de $x_0 = 0$ o $y_0 = 0$. ♦

Ejemplo 7.5 Dependencia funcional *Sean $A \subset \mathbb{R}^n$ un conjunto abierto y $f_1, \ldots, f_n : A \to \mathbb{R}$ funciones suaves. Las funciones $f_1, \ldots, f_n$ son **funcionalmente dependientes** en $x_0 \in A$ si existen una vecindad U del punto $(f_1(x_0), \ldots, f_n(x_0)) \in \mathbb{R}^n$ y una función suave $F : U \to \mathbb{R}$, tal que $\mathbf{D}F \neq 0$ en una vecindad de $(f_1(x_0), \ldots, f_n(x_0))$, y si*

$$F(f_1(x), \ldots, f_n(x)) = 0$$

para todo x en alguna vecindad de x_0.

a. *Muéstrese que si $f_1, \ldots, f_n$ son funcionalmente dependientes en x_0, entonces*

$$\frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)} = 0 \text{ en } x_0.$$

b. *Si*

$$\frac{\partial(f_1, \ldots, f_{n-1})}{\partial(x_1, \ldots, x_{n-1})} \neq 0 \quad \text{y} \quad \frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)} = 0$$

en una vecindad de x_0, muéstrese que $f_1, \ldots, f_n$ son funcionalmente dependientes y que, para alguna G,

$$f_n = G(f_1, \ldots, f_{n-1}).$$

Solución Sea $f = (f_1, \ldots, f_n)$.

a. Tenemos que $F \circ f = 0$, de modo que $\mathbf{D}F(f(x)) \circ \mathbf{D}f(x) = 0$. Si $Jf(x_0) \neq 0$, $\mathbf{D}f(x)$ sería invertible en una vecindad de x_0, lo que implica que $\mathbf{D}F(f(x)) = 0$. Por el teorema de la función inversa, esto implica que $\mathbf{D}F(y) = 0$ en toda una vecindad de $f(x_0)$.

b. Las condiciones de **b** implican que $\mathbf{D}f$ tiene rango $n - 1$. Por lo tanto, el corolario **7.4.3** implica que existen funciones g y h tales que

$$g \circ f \circ h(x_1, \ldots, x_n) = (x_1, \ldots, x_{n-1}, 0).$$

Sea F la última componente de g. Entonces

$$F(f_1, \ldots, f_n) = 0.$$

Como g es invertible, $\mathbf{D}F \neq 0$.

El teorema de la función implícita implica que $f_n = G(f_1, \ldots, f_{n-1})$; es decir, que podemos despejar $f_n = G(f_1, \ldots, f_{n-1})$ localmente en la ecuación $F(f_1, \ldots, f_n) = 0$ siempre que podamos mostrar que $\partial F/\partial y_n \neq 0$. Como vimos en la parte **a**,

$$\mathbf{D}F(f(x)) \circ \mathbf{D}f(x) = 0,$$

o bien, con la notación en componentes $y_i = f_i(x_1, \ldots, x_n)$,

$$\left(\frac{\partial F}{\partial y_1}, \ldots, \frac{\partial F}{\partial y_n}\right) \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_n}{\partial x_1} & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{pmatrix} = 0.$$

Si $\partial F/\partial y_n = 0$, tendríamos

$$\left(\frac{\partial F}{\partial y_1}, \ldots, \frac{\partial F}{\partial y_{n-1}}\right) \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_{n-1}} \\ \vdots & & \vdots \\ \dfrac{\partial f_{n-1}}{\partial x_1} & \cdots & \dfrac{\partial f_{n-1}}{\partial x_{n-1}} \end{pmatrix} = 0,$$

o

$$\left(\frac{\partial F}{\partial y_1}, \ldots, \frac{\partial F}{\partial y_{n-1}}\right) = 0,$$

pues la matriz cuadrada es invertible, por la hipótesis

$$\frac{\partial(f_1, \ldots, f_{n-1})}{\partial(x_1, \ldots, x_{n-1})} \neq 0.$$

Esto implica que $\mathbf{D}F = 0$, lo que no es cierto. Por lo tanto, $\partial F/\partial y_n \neq 0$ y obtenemos el resultado deseado.

El lector debe observar la analogía entre la dependencia lineal y la dependencia funcional, donde las condiciones sobre el rango o el determinante se reemplazan por las condiciones análogas sobre la matriz jacobiana. ♦

Ejercicios del capítulo 7

1. Escríbase una expresión para $\partial f/\partial x$ si $f(x, y) = g(x, h(x, y))$, donde $g, h : \mathbb{R}^2 \to \mathbb{R}$.

2. Considérese el siguiente conjunto de p ecuaciones con $n + p$ incógnitas:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n + a_{1,n+1}x_{n+1} + \cdots + a_{1,n+p}x_{n+p} &= 0 \\ \vdots \qquad\qquad & \\ a_{p1}x_1 + \cdots + a_{pn}x_n + a_{p,n+1}x_{n+1} + \cdots + a_{p,n+p}x_{n+p} &= 0. \end{aligned}$$

¿Qué dice el teorema de la función implícita acerca de la solución de estas ecuaciones en términos de las incógnitas $x_{n+1}, \ldots, x_{n+p}$? ¿Se reduce esto a un teorema conocido del álgebra lineal?

3. Sea $f : \mathbb{R} \to \mathbb{R}$ C^1 y

$$u = f(x), \quad v = -y + xf(x).$$

Si $f'(x_0) \neq 0$, muéstrese que esta transformación es invertible cerca de (x_0, y_0) y que la inversa tiene la forma

$$x = f^{-1}(u), \quad y = -v + uf^{-1}(u).$$

4. Muéstrese que las ecuaciones

$$x^2 - y^2 - u^3 + v^2 + 4 = 0$$
$$2xy + y^2 - 2u^2 + 3v^4 + 8 = 0$$

determinan funciones $u(x, y)$, $v(x, y)$ cerca de $x = 2$, $y = -1$ tales que $u(2, -1) = 2$, $v(2, -1) = 1$. Calcúlese $\partial u/\partial x$.

5. **a.** Defínase $x : \mathbb{R}^2 \to \mathbb{R}$ como $x(r, \theta) = r \cos\theta$ e $y : \mathbb{R}^2 \to \mathbb{R}$ como $y(r, \theta) = r$ sen θ. Muéstrese que

$$\frac{\partial(x, y)}{\partial(r, \theta)}(r_0, \theta_0) = r_0.$$

b. ¿Cuándo podemos formar una función inversa suave $r(x, y)$, $\theta(x, y)$? Verifíquese esto de forma directa y con el teorema de la función inversa.

c. Considérese la siguiente transformación para las coordenadas esféricas:

$$x(r, \varphi, \theta) = r\operatorname{sen}\varphi\cos\theta;$$
$$y(r, \varphi, \theta) = r\operatorname{sen}\varphi\operatorname{sen}\theta;$$
$$z(r, \varphi, \theta) = r\cos\theta.$$

Muéstrese que

$$\frac{\partial(x, y, z)}{\partial(r, \varphi, \theta)} = r^2\operatorname{sen}\varphi.$$

d. ¿Cuándo podemos expresar (r, φ, θ) en términos de (x, y, z)?

6. Determínese si la "curva" descrita por la ecuación $x^2 + y + \operatorname{sen}(xy) = 0$ se puede escribir en la forma $y = f(x)$ en una vecindad de $(0, 0)$. ¿Permite el teorema de la función implícita decidir si la ecuación se puede escribir en la forma $x = h(y)$ en una vecindad de $(0, 0)$?

7. Sea f una función que satisface las condiciones del teorema de la función inversa y g la inversa local $g = f^{-1} : W \to U$. Sean $x_0 \in U$ e $y_0 = f(x_0)$. Considérese el caso $n = 3$ y muéstrese que

$$Jf(x_0)\mathbf{D}_1 g_i(y_0) = \begin{vmatrix} \delta_{i,1} & \mathbf{D}_1 f_2(x_0) & \mathbf{D}_1 f_3(x_0) \\ \delta_{i,2} & \mathbf{D}_2 f_2(x_0) & \mathbf{D}_2 f_3(x_0) \\ \delta_{i,3} & \mathbf{D}_3 f_2(x_0) & \mathbf{D}_3 f_3(x_0) \end{vmatrix},$$

donde $\delta_{i,j} = 1$ si $i = j$ y 0 si $i \neq j$. Deduzca de esto la siguiente expresión para $\mathbf{D}_1 g_1$:

$$\mathbf{D}_1 g_1 = \frac{\partial(f_2, f_3)/\,\partial(x_2, x_3)}{\partial(f_1, f_2, f_3)/\,\partial(x_1, x_2, x_3)}.$$

Además, obtenga las expresiones de las otras ocho derivadas parciales $\mathbf{D}_j g_i$.

8. Sea (x_0, y_0, z_0) un punto del lugar geométrico definido por $z^2 + xy - a = 0$, $z^2 + x^2 - y^2 - b = 0$.

a. ¿En qué condiciones suficientes se puede representar la parte del lugar geométrico cercana a (x_0, y_0, z_0) en la forma $x = f(z)$, $y = g(z)$?

b. Calcúlese $f'(z)$ y $g'(z)$.

9. **a.** Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ una función suave y supóngase que

$$\frac{\partial f_1}{\partial x} = \frac{\partial f_2}{\partial y}, \quad \frac{\partial f_1}{\partial y} = -\frac{\partial f_2}{\partial x}.$$

(Éstas son las llamadas *ecuaciones de Cauchy-Riemann* y que surgen naturalmente en la teoría de variable compleja; véase, por ejemplo, J. Marsden y M. Hoffman, *Basic Complex Analysis*, 2ª ed., W.H. Freeman, Nueva York, 1987.) Muéstrese que $Jf(x, y) = 0$ sii $\mathbf{D}f(x, y) = 0$ y, por lo tanto, que f es localmente invertible sii $\mathbf{D}f(x, y) \neq 0$. Demuéstrese que la función inversa también satisface las ecuaciones de Cauchy-Riemann.

b. Muéstrese que el criterio de invertibilidad de **a** puede ser falso (dando un ejemplo) si f no satisface las ecuaciones de Cauchy-Riemann.

10. Sean f_1, f_2, f_3 funciones continuamente diferenciables de $\mathbb{R}^4$ en $\mathbb{R}$. Dense condiciones suficientes para que en las ecuaciones

$$f_1(x, y, z, t) = 0, \quad f_2(x, y, z, t) = 0, \quad f_3(x, y, z, t) = 0$$

se puedan despejar las variables x, y, z en términos de t.

11. **a.** Supóngase que $f : \mathbb{R}^n \to \mathbb{R}^m$ es de clase C^1 y que $\mathbf{D}f(x_0)$ tiene rango m. Esto significa que $\mathbf{D}f(x_0)$ es suprayectiva como transformación lineal. Muéstrese que existe toda una vecindad de $f(x_0)$ que está contenida en la imagen de f.

b. Supóngase que $f : \mathbb{R}^n \to \mathbb{R}^m$ es C^1 y que $\mathbf{D}f(x_0)$ es inyectiva. Muéstrese que f es inyectiva en una vecindad de x_0.

12. Muéstrese que el teorema de la función implícita implica al teorema de la función inversa.

13. Demuéstrese el corolario **7.2.2.**

14. Si $f_1, \ldots, f_k$ son funcionalmente independientes en $\mathbb{R}^n$ (es decir, si $\mathbf{D}f$ tiene rango k para $f = (f_1, \ldots, f_k), k \leq n$) y $g, f_1, \ldots, f_k$ son funcionalmente dependientes, muéstrese que localmente podemos escribir $g = F(f_1, \ldots, f_k)$.

15. Considérese la transformación $\mathcal{L}^{-1} : GL(\mathbb{R}^n, \mathbb{R}^n) \to GL(\mathbb{R}^n, \mathbb{R}^n)$, $A \mapsto A^{-1}$, que aplica cada matriz en su inversa. Muéstrese que la diferencial de esta transformación está dada por

$$\mathbf{D}\mathcal{L}^{-1}(A) \cdot B = -A^{-1} \circ B \circ A^{-1}.$$

16. Proporciónese una demostración directa del lema de Morse para funciones $f : \mathbb{R} \to \mathbb{R}$. ¿Se aplica ésta a

a. $f(x) = x^3$?

b. $f(x) = x\operatorname{sen}(1/x)$?

17. ¿Debe ser única la función h en el teorema de rectificación del dominio? Razónese la respuesta.

18. Demuéstrese la siguiente generalización del teorema de rectificación del dominio: Sean $A \subset \mathbb{R}^n$ un conjunto abierto y $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$, una función de clase C^p. Sea $x_0 \in A$ y supóngase que $f(x_0) = 0$ y rango $\mathbf{D}f(x_0) = m$. Entonces existen un conjunto abierto U, un conjunto abierto V que contiene a x_0 y una función $h : U \to V$ de clase C^p, con inversa $h^{-1} : V \to U$ de clase C^p (es decir, un difeomorfismo C^p) tales que $f(h(x_1, \ldots, x_n)) = (x_{n-m+1}, \ldots, x_n)$. [Sugerencia: si $\mathbf{D}f(x_0)$ tiene rango m, deben existir $j_i, \ldots, j_m$ tales que la matriz $(\mathbf{D}_{j_k}f_i)$, $1 \leq i, k \leq m$ es invertible. Defínase la aplicación de permutación $g : \mathbb{R}^n \to \mathbb{R}^n$ como $g(x_1, \ldots, x_n) = (x_1, \ldots, x_{j_1-1}, x_{n-m+1}, x_{j_1+1}, \ldots, x_{jm-1}, x_n, x_{jm+1}, \ldots, x_{n-m}, x_{j_1}, \ldots, x_{jm})$.]

19. Sea $F : \mathbb{R}^4 \to \mathbb{R}^2$ dada por $F(x, y, u, v) = (u^3 + vx + y, uy + v^3 - x)$. ¿En qué puntos podemos expresar u, v en términos de x, y en la ecuación $F(x, y, u, v) = 0$? Calcúlese $\partial u/\partial x$.

20. "Si $f(x, y, z) = 0$, entonces $\partial x/\partial y \cdot \partial y/\partial x \cdot \partial x/\partial z = -1$." ¿Qué piensa el lector que significa esto en realidad? (Los libros de termodinámica son notables por este tipo de afirmaciones tan misteriosas.)

21. Sean $f : \mathbb{R}^n \to \mathbb{R}^n$ y $g_i : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, n$ funciones de clase C^1. Defínase $h : \mathbb{R}^n \to \mathbb{R}^n$ como $h(x) = f(g_1(x_1), \ldots, g_n(x_n))$, donde $g = (g_1, \ldots, g_n)$ y $x = (x_1, \ldots, x_n)$. Muéstrese que

$$\mathbf{D}h(x) = \mathbf{D}f(g_1(x_1), \ldots, g_n(x_n)) \begin{pmatrix} g_1'(x_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & g_n'(x_n) \end{pmatrix}$$

22. Sea $f(x, y) = (xy(x^2 - y^2))/(x^2 + y^2)$ para $(x, y) \neq (0, 0)$ y $f(0, 0) = 0$. ¿Es f de clase C^2?

23. Sea $C \subset \mathbb{R}^n$ un subconjunto cerrado tal que $x \in C \Rightarrow \alpha x \in C$ para todo $\alpha \geq 0$.

a. Analícese la "forma" de C.

b. Sea $f : C \to \mathbb{R}^n$ continua y $f(\alpha x) = \alpha f(x)$ para $x \in C$, $\alpha \geq 0$. Muéstrese que existe M tal que $\|f(x)\| \leq M\|x\|$ para todo $x \in C$.

24. Sean $f(x, y, z) = x^2 - yz - \operatorname{sen}(xz)$ y

$$g(x, y, z) = (x \cos y, x \operatorname{sen} y \cos z, x \operatorname{sen} y \operatorname{sen} z).$$

Calcúlese la derivada de $f \circ g$.

25. Sean $B(0, r) = \{x \in \mathbb{R}^n \mid \|x\| \leq r\}$ y $f : B(0, r) \to \mathbb{R}^n$ una aplicación tal que

a. $\|f(x) - f(y)\| \leq \frac{1}{3}\|x - y\|$

b. $\|f(0)\| \leq \frac{2}{3}r$

Demuéstrese que existe un único $x \in B(0, r)$ tal que $f(x) = x$.

26. Muéstrese que existen números positivos $p > 0$, $q > 0$ y funciones únicas u, v de $]-1-p, -1+p[$ en $]1-q, 1+q[$ para las que

$$xe^{u(x)} + u(x)e^{v(x)} = 0 = xe^{v(x)} + v(x)e^{u(x)}$$

para todo $x \in\,]-1-p, -1+p[$ y $u(-1) = 1 = v(-1)$.

27. Obténgase una estimación del intervalo de tiempo en que existe la solución de $dx/dt = t^2x^3e^{tx}$, $x(0) = 1$.

28. Sea $A \subset \mathbb{R}^n$ un conjunto compacto y $B \subset \mathcal{C}(A, \mathbb{R})$ uno compacto. Muéstrese que existen $f_0 \in B$ y $x_0 \in A$ tales que $g(x) \leq f_0(x_0)$ para todo $g \in B$ y $x \in A$.

29. Sean $a_n \geq a_{n+1} \geq 0$ y $a_n \to 0$. Sea $f(x) = \sum_{n=0}^{\infty} a_n x^n$. Muéstrese que $f(x)$ es continua en $[-1, 0]$.

30. ¿Es posible despejar $u(x, y, z)$, $v(x, y, z)$ en las ecuaciones

$$xy^2 + xzu + yv^2 = 3$$
$$u^3yz + 2xv - u^2v^2 = 2$$

cerca de $(x, y, z) = (1, 1, 1)$, $(u, v) = (1, 1)$? Calcúlese $\partial v/\partial y$.

31. Considérese la ecuación $dx/dt = 1 + tx$, $x(0) = 0$. Examínese el esquema de iteración dado en **§7.5** para obtener una expresión de la solución en serie de potencias. Examínese el radio de convergencia.

32. Calcúlese el índice de la función $x^2 + y^2 - 7x - 8y + xy + 16 + (x - 2)^3$ en su punto crítico $x = 2$, $y = 3$. Analícese la naturaleza de la función cerca de este punto.

33. Proporciónese otra demostración del lema de Morse, como sigue. Supóngase que $x_0 = 0$ y que $f(x_0) = 0$. Úsese el teorema de Taylor para escribir $f(x) = \frac{1}{2}\mathbf{D}^2 f(0)\cdot(x, x) + \frac{1}{2}R_x(x, x) = \frac{1}{2}\langle A_x x, x\rangle$, de modo que para cada x, A_x sea una transformación lineal simétrica de $\mathbb{R}^n$. Por hipótesis, A_0 es un isomorfismo, de modo que A_x es un isomorfismo si x está cerca de 0. Sea $Q_x = A_0 A_x^{-1}$, de modo que $Q_0 = I$. Usando una serie de potencias, podemos definir la raíz cuadrada T_x de Q_x para x cercano a 0; es decir, $T_x^2 = Q_x$. Muéstrese que $Q_x A_x = A_x Q_x^T$, donde T indica la matriz transpuesta, y úsese la serie de potencias de T_x para mostrar que la misma ecuación es válida para T_x. Sea $S_x = T_x^{-1}$ y concluyase que $A_x = S_x A_0 S_x^T$. Sea $h(x) = S_x^T x$ y muéstrese que $\mathbf{D}h(0) = I$; aplíquese ahora el teorema de la función inversa para concluir que h es localmente invertible. Sea $g = h^{-1}$. Muéstrese que

$$f(x) = \frac{1}{2}\langle A_0 h(x), h(x)\rangle$$

y dedúzcase que $f \circ g(x) = \frac{1}{2}\mathbf{D}^2 f(0)(x, x)$. Por último, úsese un cambio lineal de coordenadas para diagonalizar la forma cuadrática $\frac{1}{2}\mathbf{D}^2 f$.

Encuéntrense los extremos relativos de $f|S$ en los ejercicios del 34 al 37.

34. $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x^2 + y^2$, $\quad S = \{(x, 2) \mid x \in \mathbb{R}\}$.

35. $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x^2 + y^2$, $S = \{(x, y) \mid x^2 - y^2 = 1\}$.

36. $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x^2 - y^2$, $S = \{(x, \cos x) \mid x \in \mathbb{R}\}$.

37. $f : \mathbb{R}^3 \to \mathbb{R}$, $(x, y, z) \mapsto x^2 + y^2 + z^2$, $S = \{(x, y, z) \mid z \geq -2 + x^2 + y^2\}$.

38. Una caja rectangular, sin tapa, debe tener un área superficial de 16 metros cuadrados. Determínense las dimensiones que maximizan su volumen.

39. Diséñese un recipiente cilíndrico para un litro de agua que use la mínima cantidad de metal.

Capítulo 8
Integración

Sin duda el lector está familiarizado con el proceso de integración para funciones de una variable y su aplicación a problemas prácticos relacionados con áreas, volúmenes, longitudes de arco, etcétera. El propósito de este capítulo y del siguiente es revisar, consolidar y ampliar este conocimiento. En este capítulo formulamos las definiciones básicas para una teoría general de integración. Es útil, aunque no esencial, cierta familiaridad con las situaciones sencillas relacionadas con las integrales múltiples.

Los poderosos teoremas de cálculo de integrales múltiples aparecen en el siguiente capítulo. Éstos son el *teorema de Fubini*, que nos permite reducir una integral múltiple a integrales simples iteradas y la *fórmula del cambio de variables*, que nos permite pasar de las coordenadas rectangulares a otro sistema de coordenadas más conveniente, como las coordenadas polares o esféricas y, más en general, cambiar de un sistema de coordenadas a otro. Para obtener una teoría satisfactoria de las integrales múltiples, aun para funciones continuas, es conveniente introducir el concepto de conjunto de "medida nula". Veremos que uno de los principales teoremas establece que una función acotada es integrable sii sus discontinuidades forman un conjunto de medida nula. En particular, una función acotada con una cantidad finita o numerable de discontinuidades es integrable.

§8.1 Funciones integrables

Basándonos en el trabajo que hicimos en **§4.8** vamos a bosquejar la forma en que debería desarrollarse la integración en $\mathbb{R}^2$. Supóngase que $f\colon A \subset \mathbb{R}^2 \to \mathbb{R}$ es una función acotada no negativa (véase la figura 8.1-1), donde A es un conjunto acotado. La gráfica de la función f es una superficie en $\mathbb{R}^3$, y el proceso de integración se usa para determinar el volumen bajo esta superficie. Encerramos A en algún rectángulo $[a_1, b_1] \times [a_2, b_2]$ y extendemos f a todo el rectángulo, definiéndola como cero fuera de A. Después dividimos $[a_1, b_1] \times [a_2, b_2]$ en rectángulos más pequeños, estableciendo una partición de

$[a_1, b_1]$, por ejemplo, $a_1 = x_0 < x_1 < \cdots < x_{n-1} < x_n = b_1$, y una partición de $[a_2, b_2]$, por ejemplo, $a_2 = y_0 < y_1 < \cdots < y_{m-1} < y_m = b_2$, formando así mn rectángulos $[x_i, x_{i+1}] \times [y_j, y_{j+1}]$. Después calculamos el volumen del bloque sombreado de la figura 8.1-1 como $\inf\{f(z) \mid z \in [x_i, x_{i+1}] \times [y_j, y_{j+1}]\} \cdot (x_{i+1} - x_i)(y_{j+1} - y_j)$ y el volumen del bloque sombreado más el bloque marcado con diagonales de la figura 8.1-1 como

$$\sup\{f(z) \mid z \in [x_i, x_{i+1}] \times [y_j, y_{j+1}]\} \cdot (x_{i+1} - x_i)(y_{j+1} - y_j).$$

Sumamos estos volúmenes para todo i y todo j (es decir, los mn "subrectángulos" del rectángulo $[a_1, b_1] \times [a_2, b_2]$) para obtener dos valores, denotados $L(f, P)$ y $U(f, P)$, donde P representa la partición. Si

$$\sup\{L(f, P) \mid P \text{ es una partición}\} = \inf\{U(f, P) \mid P \text{ es una partición}\},$$

decimos que f es ***integrable (Riemann)*** sobre A y definimos la ***integral (de Riemann)*** sobre el conjunto A, denotada $\int_A f$ o $\iint_A f(x, y)\, dx\, dy$, como

$$\int_A f = \sup\{L(f, P)\} = \inf\{U(f, P)\}.$$

Ya hemos presentado la mayoría de las ideas necesarias para una teoría de la integración de funciones acotadas sobre conjuntos acotados de dimensiones arbitrarias. Lo que resta es formalizar las afirmaciones para $\mathbb{R}^n$.

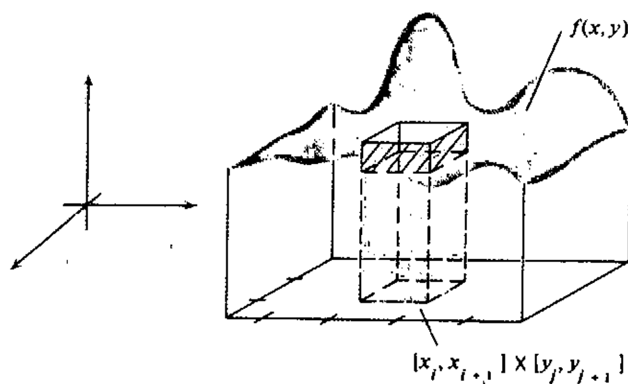


FIGURA 8.1-1 La gráfica de una función acotada y una columna que ayuda a conformar el volumen por debajo de ella

Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$ una función acotada cuyo dominio es un conjunto acotado A. Elíjase un rectángulo $[a_1, b_1] \times \cdots \times [a_n, b_n]$ que encierre A. Además, definamos f sobre

todo el rectángulo, haciéndola igual a 0 en los puntos no contenidos en A. Sea P una partición de $[a_1, b_1] \times \cdots \times [a_n, b_n]$ obtenida al dividir cada $[a_i, b_i]$ mediante los puntos $x_0^i, \ldots, x_{m_i}^i$ y formando los $m_1 m_2 \cdots m_n$ ***rectángulos***

$$[x_{j_1}^1, x_{j_1+1}^1] \times \cdots \times [x_{jn}^n, x_{jn+1}^n], \quad \text{donde} \quad 0 \leq j_i \leq m_i - 1.$$

Definimos el ***volumen*** del rectángulo $B = [a_1, b_1] \times \cdots \times [a_n, b_n]$ como el producto de las longitudes de las aristas: $\text{vol}(B) = v(B) = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n)$. Sea $L(f, P)$ la ***suma inferior*** de f para P, definida como

$$L(f,P) = \sum_{R\in P} [\inf\{f(x) \mid x \in R\}]v(R),$$

donde la suma se realiza sobre todos los subrectángulos R de la partición P. Análogamente, sea $U(f, P)$ la ***suma superior*** de f para P, definida por

$$U(f,P) = \sum_{R\in P} [\sup\{f(x) \mid x \in R\}]v(R).$$

Observemos ahora algunas propiedades de $L(f, P)$ y $U(f, P)$. Por la definición, vemos que para cada partición P, $L(f, P) \leq U(f, P)$. Supóngase ahora que P' es cualquier partición que sea un ***refinamiento*** de o que sea ***más fina que*** P; esto significa que cada subrectángulo perteneciente a P' está *contenido* en un subrectángulo perteneciente a P. Afirmamos que $L(f, P) \leq L(f, P')$. De hecho, esto es consecuencia de que el ínfimo de f en un rectángulo es menor o igual que el ínfimo en cualquier rectángulo contenido en él. Análogamente, $U(f, P') \leq U(f, P)$. Esto tiene la siguiente consecuencia: *si* P' *y* P'' *son dos particiones cualesquiera de* $[a_1, b_1] \times \cdots \times [a_n, b_n]$, *entonces* $L(f, P') \leq U(f, P'')$. Para demostrarlo, sea P una partición del rectángulo que refine P' y P'', lo que podemos lograr si usamos todos los puntos de subdivisión de P' y P''; entonces, $L(f, P') \leq L(f, P) \leq U(f, P) \leq U(f, P'')$.

Como la función f y el rectángulo B están acotados, los conjuntos de números $\{L(f, P) \mid P \text{ es una partición de } B\}$ y $\{U(f, P) \mid P \text{ es una partición de } B\}$ están acotados inferiormente por $\inf\{f(x) \mid x \in B\}\, \text{vol}(B)$ y superiormente por $\sup\{f(x) \mid x \in B\}\, \text{vol}(B)$. Cada número del primer conjunto es menor o igual que todos los números del segundo conjunto. Así, si denotamos

$$s = \sup\{L(f, P) \mid P \text{ es una partición de } B\}$$
$$S = \inf\{U(f, P) \mid P \text{ es una partición de } B\}$$

entonces $s \leq S$. Con esta notación, podemos hacer otra definición.

8.1.1 Definición *Si f es una función acotada en una región acotada A, entonces la* ***integral superior de f sobre A*** *se define como*

$$\overline{\int_A} f = S = \inf\{U(f,P) \mid P \text{ es una partición de } B\}$$

y la ***integral inferior de f sobre A*** *como*

$$\underline{\int_A} f = s = \sup\{L(f,P) \mid P \text{ es una partición de } B\}$$

donde B es cualquier rectángulo que contenga la región A. Decimos que f es ***integrable Riemann*** *(de ahora en adelante sólo usaremos la palabra "integrable") si $s = S$ y definimos la* ***integral de f sobre A*** *como su valor común:*

$$\int_A f = \overline{\int_A} f = \underline{\int_A} f$$

En vez de $\int_A f$, con frecuencia se usa la notación $\int_A f(x)\, dx$ o $\int \cdots \int_A f(x_1, \ldots, x_n)\, dx_1 \cdots dx_n$. Si $f: [a, b] \to \mathbb{R}$, también se usa la notación $\int_a^b f$ o $\int_a^b f(x)\, dx$.

Existe una importante caracterización equivalente de la integral de Riemann, que presentamos en el siguiente teorema.

8.1.2 Teorema de Darboux *Sea $A \subset \mathbb{R}^n$ acotado y contenido en algún rectángulo S. Sea $f: A \to \mathbb{R}$ acotada y extendámosla a S definiendo $f = 0$ fuera de A. Entonces f es integrable con integral I sii para todo $\varepsilon > 0$ existe $\delta > 0$ tal que si P es cualquier partición de S en rectángulos $S_1, \ldots, S_N$, con lados de longitud $< \delta$ y si $x_1 \in S_1, \ldots, x_N \in S_N$, tenemos*

$$\left| \sum_{i=1}^{N} f(x_i) v(S_i) - I \right| < \varepsilon.$$

Decimos que $\sum_{i=1}^{N} F(x_i) v(S_i)$ es una ***suma de Riemann***.

Una condición muy relacionada con el teorema **8.1.2** es l siguiente.

8.1.3 Condición de Riemann *f es integrable sii ıra todo $\varepsilon > 0$ existe una partición P_ε de S tal que $0 \le U(f, P_\varepsilon) - L(f, P_\varepsilon) < \varepsilon$.*

La demostración de la condición de Riemmann se dará junto con la demostración del teorema de Darboux al final del capítulo.

Obsérvese que si f es continua, podemos interpretar la sumas superiores e inferiores para una partición como sumas de Riemann especiales, pues f alcanza su máximo y su mínimo en cada subrectángulo en algún punto de tal subrectángulo. Si f es continua en todo el rectángulo S (que es un intervalo si $n = 1$), entonces la continuidad uniforme de f (véase §**4.6**) y la condición de Riemann implican que f es integrable. De hecho, demostraremos un resultado más general en **8.3.1.**

Ahora vamos a ilustrar algunas de estas ideas en el caso particular de una dimensión. Esto ya se hizo en §**4.8**, pero merece la pena darles un repaso en este momento.

8.1.4 Ejemplo *Interprétense las sumas de Riemann geométricamente, para* $f: [a, b] \to \mathbb{R}$.

Solución Sea $P : a = x_0 < x_1 < \cdots < x_n = b$ una partición y $c_i \in [x_i, x_{i+1}]$. Por definición, la suma de Riemann es

$$R = \sum_{i=0}^{n-1} f(c_i)(x_{i+1} - x_i),$$

que es el área total de los rectángulos representados en la figura 8.1-2, donde el área de los rectángulos que están por debajo del eje x se cuenta con un signo negativo. Observemos que $L(f, P) \leq R \leq U(f, P)$, de modo que el teorema de Darboux es plausible. ♦

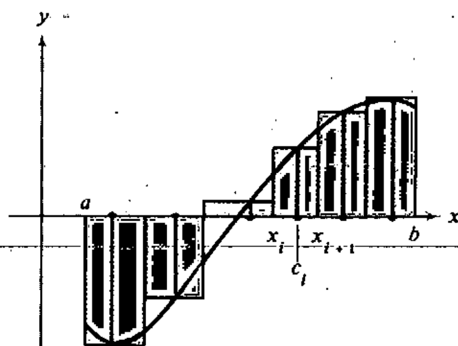


FIGURA 8.1-2 Aproximación de una integral mediante una suma de Riemann

8.1.5 Ejemplo *Muéstrese que $\int_0^1 x\, dx = 1/2$, usando la definición de la integral. Compárese con un cálculo geométrico.*

Solución Separamos [0, 1] en n partes iguales

$$\left[0, \frac{1}{n}\right], \left[\frac{1}{n}, \frac{2}{n}\right], \ldots, \left[\frac{n-1}{n}, 1\right].$$

Usando esto como partición, el ínfimo de $f(x) = x$ en $[i/n, (i+1)/n]$ es i/n y el supremo es $(i+1)/n$. Así, llamando P a esta partición,

$$U(f, P) = \sum_{i=0}^{n-1} \frac{i+1}{n} \cdot \frac{1}{n} = \frac{1}{n^2} \cdot \sum_{i=0}^{n-1} (i+1)$$
$$= \frac{1}{n^2}(1 + 2 + \cdots + n) = \frac{1}{n^2} \cdot \frac{1}{2} n \cdot (n+1)$$

y

$$L(f, P) = \sum_{i=0}^{n-1} \frac{i}{n} \cdot \frac{1}{n} = \frac{1}{n^2}(0 + 1 + \cdots + (n-1)) = \frac{1}{n^2}\left(\frac{1}{2}\right)(n-1)(n)$$

puesto que $1 + 2 + \cdots + k = (1/2)k(k+1)$. Así,

$$U(f, P) = \frac{1}{2}\left(1 + \frac{1}{n}\right) \quad \text{y} \quad L(f, P) = \frac{1}{2}\left(1 - \frac{1}{n}\right).$$

Ambas convergen a 1/2 cuando $n \to \infty$. Así, por la condición de Riemann (o el teorema de Darboux), vemos que f es integrable, y su integral es igual a 1/2. Esto también está claro geométricamente, en la figura 8.1-3. ♦

Ejercicios de §8.1

1. Demuéstrese que si R es una suma de Riemann para una función f y una partición P, entonces $L(f, P) \leq R \leq U(f, P)$.

2. Sea $f: [0, 2] \to \mathbb{R}$ dada por $f(x) = 0$ si $0 \leq x \leq 1$, y $f(x) = 1$ si $1 < x \leq 2$. Calcúlese $\int_0^2 f(x)\, dx$ mediante la definición.

3. Sea $f: [0, 1] \to \mathbb{R}$ dada por $f(x) = 0$ si $x \neq 1/2$ y $f(1/2) = 1$. Demuéstrese que f es integrable y que $\int_0^1 f(x)\, dx = 0$.

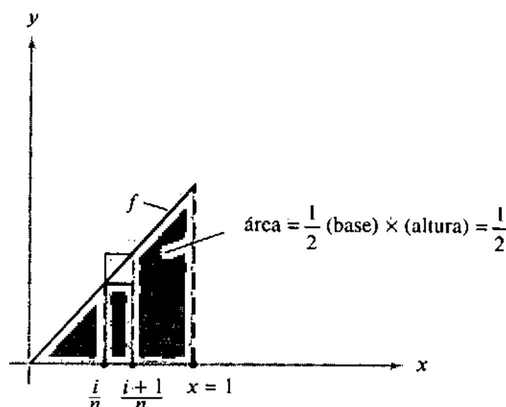


FIGURA 8.1-3 La integral de $f(x) = x$ es el área de un triángulo

4. Sean $A \subset \mathbb{R}^n$ y $f(x) = 1$ para $x \in A$. ¿Cuál cree el lector que deba ser el valor de $\int_A f$?

5. Evalúese $\int_0^1 (3x + 4)dx$ usando la definición, y compárese la respuesta con un cálculo geométrico del área.

6. Sea $f : [a, b] \to \mathbb{R}$ continua. Úsese la condición de Riemann y la continuidad uniforme de f para demostrar que f es integrable.

§8.2 Volumen y conjuntos de medida nula

En la recta real, integramos generalmente sobre intervalos. Sin embargo, en $\mathbb{R}^n$ necesitamos por lo general integrar sobre conjuntos más complicados. Debemos asegurarnos de que los conjuntos con los que trabajamos estén restringidos de modo que la partición en la definición de integrabilidad sea razonable. En este caso, "razonable" quiere decir, *grosso modo*, que la frontera del conjunto no sea demasiado complicada. Nuestro objetivo inmediato será el desarrollo de los conceptos suficientes para precisar estas ideas. En primer lugar, definimos el volumen de un conjunto.

8.2.1 Definición *Si $A \subset \mathbb{R}^n$ es un conjunto acotado, la **función característica** 1_A de A es la aplicación $1_A : \mathbb{R}^n \to \mathbb{R}$ definida por $1_A(x) = 1$ si $x \in A$ y $1_A(x) = 0$ si $x \notin A$. Decimos que A **tiene volumen** si 1_A es integrable, y el **volumen** de A es el número*

$$\int_A 1_A(x)\, dx = v(A).$$

Esta definición es natural, pues la región debajo de la gráfica de 1_A es la región "cilíndrica" con altura 1 y base A (figura 8.2-1). También usaremos la frase "A tiene contenido" para indicar lo mismo que "A tiene volumen". A veces, un conjunto que tiene volumen se llama ***medible Jordan***.

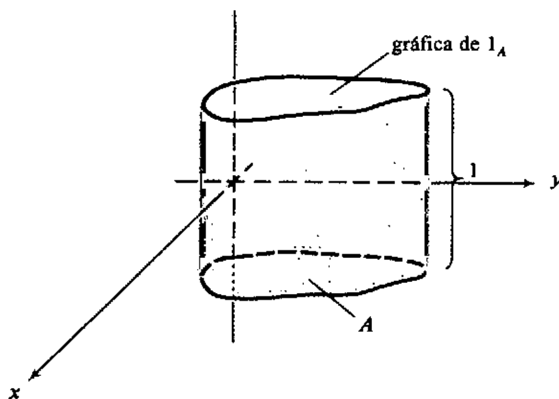


FIGURA 8.2-1 La función característica de un conjunto

Si $n = 1$, de modo que $A \subset \mathbb{R}$, decimos que $v(A)$ es la ***longitud*** de A y cuando $A \subset \mathbb{R}^2$, usamos el término ***área*** de A para $v(A)$.

Decimos que A tiene ***volumen nulo*** (o ***contenido nulo***) si $v(A) = 0$. Por la definición de integral, esto es equivalente a la afirmación de que para cada $\varepsilon > 0$ existe un recubrimiento finito de A mediante rectángulos, digamos $S_1, \ldots, S_m$, de modo que el volumen usual sea menor que ε; es decir,

$$\sum_{i=1}^{m} v(S_i) < \varepsilon,$$

donde $v(S_i)$ para rectángulos se calcula de la forma usual (los detalles aparecen en el ejemplo resuelto **8.1** al final del capítulo).

Es útil permitir el uso de recubrimientos *numerables* al igual que finitos. Estas ideas fueron presentadas de forma sistemática por Henri Lebesgue hacia el año 1900.

8.2.2 Definición *Un conjunto $A \subset \mathbb{R}^n$ (no necesariamente acotado) tiene **medida nula** si para cada $\varepsilon > 0$ existe un recubrimiento de A, digamos $S_1, S_2, \ldots$, mediante una cantidad numerable (o finita) de rectángulos tales que el volumen total $\sum_{i=1}^{\infty} v(S_i) < \varepsilon$. Recuérdese que $S_1, S_2, \ldots$ **recubren** A si $\bigcup_{i=1}^{\infty} S_i \supset A$. Obsérvese que estos rectángulos se pueden solapar.*

Es importante notar que estos conceptos dependen del espacio donde se trabaje, lo que ilustramos mediante un ejemplo.

8.2.3 Ejemplo *Muéstrese que, considerado como un subconjunto de $\mathbb{R}^2$, la recta real tiene medida nula, pero como subconjunto de $\mathbb{R}$ no.*

Solución Para demostrar la primera afirmación, dado $\varepsilon > 0$, queremos encontrar rectángulos $S_1, S_2, \ldots$ que encierren el eje X y que tengan área total $< \varepsilon$. Considerando la figura 8.2-2, sean

$$S_i = [-i, i] \times \left[-\frac{\varepsilon}{2i \cdot 2^{i+1}}, \frac{\varepsilon}{2i \cdot 2^{i+1}}\right].$$

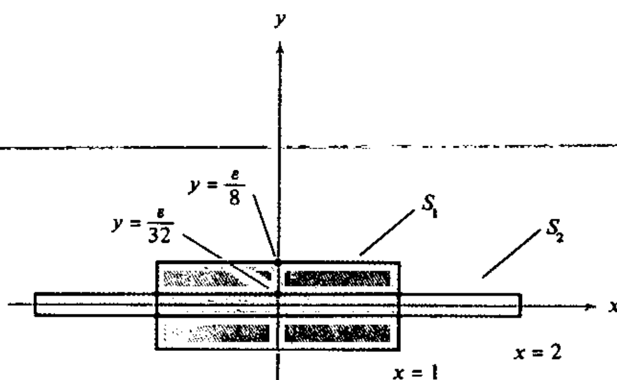


FIGURA 8.2-2 Una familia de rectángulos que recubren el eje horizontal

Como $v(S_i) = 2i[(2\varepsilon)/(2i \cdot 2^{i+1})] = \varepsilon/2^i$, obtenemos $\sum_{i=1}^{\infty} v(S_i) < \sum_{i=1}^{\infty} (\varepsilon/2^i) = \varepsilon$, ya que $1/2 + 1/4 + 1/8 + \cdots = 1$. Está claro que la recta real, como subconjunto de sí misma, no puede tener medida nula, pues al recubrirla mediante intervalos, la longitud total de los mismos será $+\infty$. ♦

Esta demostración es típica de la forma en que se demuestra que un conjunto tiene medida nula. Otro ejemplo de conjunto de medida nula es la esfera

$$S = \{x \in \mathbb{R}^n \mid \|x\| = 1\}.$$

Por la definición de volumen, está claro que *si A tiene volumen nulo, entonces A tiene medida nula*. De hecho, si A tiene volumen nulo y $\varepsilon > 0$, podemos incluso encon-

trar un recubrimiento *finito* para A con volumen total $< \varepsilon$. Además, obsérvese que si A tiene medida nula y $B \subset A$, entonces B también tiene medida nula.

La principal ventaja de la medida nula sobre el volumen nulo aparece en el siguiente teorema.

8.2.4 Teorema *Supóngase que los conjuntos $A_1, A_2, \ldots$ tienen medida nula en $\mathbb{R}^n$. Entonces $A_1 \cup A_2 \cup \cdots$ tiene medida nula en $\mathbb{R}^n$.*

De esto podemos concluir, por ejemplo, que cualquier conjunto compuesto por una cantidad numerable de puntos tiene medida nula.

8.2.5 Ejemplo Considérese el conjunto A de racionales en $[0, 1] \subset \mathbb{R}$. El conjunto A no tiene volumen; es decir, 1_A no es integrable. De hecho, la función que tiene el valor 1 para los racionales y 0 para los irracionales, no es integrable, como vimos en §**4.8.** Sin embargo, el conjunto A tiene medida nula, pues un punto tiene volumen y medida nulos y A consta de una cantidad numerable de puntos, por lo que podemos aplicar el teorema **8.2.4**. ♦

Ejercicios de §8.2

1. Muéstrese que $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ tiene volumen nulo.

2. Muéstrese que el plano XY en $\mathbb{R}^3$ tiene medida tridimensional nula.

3. Si $A \subset [a, b]$ tiene medida nula en $\mathbb{R}$, demuéstrese que $[a, b]\backslash A$ *no* tiene medida nula en $\mathbb{R}$ (el ejercicio 10 incluido al final del capítulo implica que $[a, b]$ no tiene medida nula).

4. Úsese el ejercicio 3 para mostrar que los irracionales en $[0,1]$ no tienen medida nula.

5. ¿Debe tener medida nula la frontera de un conjunto?

6. ¿Debe tener medida nula la frontera de un conjunto de medida nula?

§8.3 Teorema de Lebesgue

Consideremos ahora uno de los resultados más importantes en la teoría de la integración. Intuitivamente pensamos que la mayoría de las funciones "decentes", como las

funciones continuas, deben ser integrables, pues el área por debajo de sus gráficas debe poderse definir. Para dejar en claro de manera precisa la cuestión de cuán decente es "decente" tenemos el teorema de H. Lebesgue. Con este teorema, Lebesgue abrió nuevos avances en la teoría de integración enfatizando el concepto de medida nula.

8.3.1 Teorema de Lebesgue *Sean $A \subset \mathbb{R}^n$ un conjunto acotado y $f: A \to \mathbb{R}$ una función acotada. Extiéndase f a todo $\mathbb{R}^n$ definiéndola como cero en los puntos no contenidos en A. Entonces f es integrable (Riemann) sii los puntos donde la función f es discontinua forman un conjunto de medida nula.*

Podemos extraer dos conclusiones importantes de este resultado, como se establece en los siguientes dos corolarios.

8.3.2 Corolario *Un conjunto acotado $A \subset \mathbb{R}^n$ tiene volumen sii la frontera de A tiene medida nula.*

Esto es consecuencia de la aplicación del teorema a la función característica 1_A. Como los conjuntos finitos y numerables tienen medida nula, el teorema también implica:

8.3.3 Corolario *Sea $A \subset \mathbb{R}^n$ acotado y con volumen. Una función acotada $f: A \to \mathbb{R}$ con una cantidad finita o numerable de puntos de discontinuidad es integrable.*

Este resultado muestra que la mayoría de las funciones comunes que se encuentran en la práctica son integrables. Por ejemplo, una función continua en un intervalo $[a, b]$ es integrable, pues $[a, b]$ tiene volumen (la frontera consta de dos puntos). Las funciones con un número finito de puntos de discontinuidad también son integrables.

Obsérvese que la integrabilidad de f en el teorema **8.3.1** depende de la extensión de f. Por ejemplo, si A es el conjunto de racionales en $[0,1]$ y f es idénticamente 1, f restringida a A es continua en A pero la función f extendida no es continua en ningún punto y, de hecho, no es integrable. En el corolario **8.3.3**, *no* es necesario extender f, pues se usa el hecho de que A tiene volumen y el corolario **8.3.2**. Otro resultado útil es el siguiente.

8.3.4 Teorema

i. *Sea $A \subset \mathbb{R}^n$ un conjunto acotado, de medida nula y $f: A \to \mathbb{R}$ cualquier función integrable (acotada). Entonces $\int_A f(x)\, dx = 0$.*

ii. *Si $f : A \to \mathbb{R}$ es integrable y $f(x) \geq 0$ para todo x y $\int_A f(x)\, dx = 0$, entonces el conjunto $\{x \in A \mid f(x) \neq 0\}$ tiene medida nula.*

Este teorema es razonable, pues un conjunto de medida nula es "pequeño" y, esencialmente, tiene volumen nulo, de modo que la integral de cualquier función sobre él debería anularse. La segunda parte también es razonable por el mismo motivo.

8.3.5 Ejemplo *Sea*

$$f(x) = \begin{cases} x, & -1 \leq x \leq 0 \\ 3x + 8, & 0 < x \leq 1. \end{cases}$$

Muéstrese que f es integrable en $[-1, 1]$.

Solución El conjunto $[-1, 1]$ tiene volumen y f solamente tiene una discontinuidad, en $x = 0$. Como f está acotada, **8.3.3** implica que f es integrable. ♦

8.3.6 Ejemplo *Sea* $f(x) = \text{sen}(1/x)$ *para* $x \neq 0$ *y* $f(0) = 0$. *Muéstrese que f es integrable en* $[-1, 1]$.

Solución En este caso, f tiene un punto de discontinuidad, en $x = 0$. Además, $|f(x)| \leq 1$, de modo que f está acotada. Así, por el corolario **8.3.3**, f es integrable. ♦

8.3.7 Ejemplo Sea $f(x, y) = x^2 + \text{sen}(1/y)$ para $y \neq 0$ y $f(x, 0) = x^2$. Muéstrese que f es integrable en $A = \{(x, y) \mid x^2 + y^2 < 1\}$.

Solución *En este caso, f está acotada en A, el interior del disco unidad en* $\mathbb{R}^2$, *y tiene discontinuidades sobre la recta* $y = 0$, *que es un conjunto de medida nula en* $\mathbb{R}^2$. *Además, A tiene volumen (su frontera tiene medida nula). Por lo tanto, por el teorema* **8.3.1**, *f es integrable.* ♦

Ejercicios de §8.3

1. Sea $f(x) = x^3$ en $[-1, 1]$. Demuéstrese que f es integrable.

2. Sea $f(x, y) = 1$ si $x \neq 0$ y $f(0, y) = 0$. Demuéstrese que f es integrable en $A = [0, 1] \times [0, 1] \subset \mathbb{R}^2$.

3. Calcúlese $\int_A f$ donde f y A son como en el ejercicio 2.

4. Sea $A \subset \mathbb{R}^n$ un conjunto abierto con volumen y sean $f: A \to \mathbb{R}$ continua, $f(x) \geq 0$ y $f(x_0) > 0$ para algún $x_0 \in A$. Muéstrese que $\int_A f > 0$.

5. Sea $r_k = 1/k$, $k = 1, 2, \ldots$ y sea U el siguiente conjunto abierto en $\mathbb{R}$:

$$U = \bigcup_{k=1}^{\infty} D\left(r_k, \frac{1}{2^k}\right).$$

Decídase si U tiene volumen o no.

6. Sea $f(x) = \cos(1/x)$ para $x \neq 0$ y $f(0) = 0$. Muéstrese que f es integrable en $[-1, 1]$.

§8.4 Propiedades de la integral

Ahora presentamos algunas de las propiedades elementales de la integral, análogas a las propiedades de las funciones definidas en intervalos.

8.4.1 Teorema *Sean A, B subconjuntos acotados de $\mathbb{R}^n$, $c \in \mathbb{R}$ y $f, g : A \to \mathbb{R}$ funciones integrables. Entonces*

i. *$f + g$ es integrable y $\int_A (f + g) = \int_A f + \int_A g$.*

ii. *cf es integrable y $\int_A (cf) = c \int_A f$.*

iii. *$|f|$ es integrable y $\left|\int_A f\right| \leq \int_A |f|$.*

iv. *Si $f \leq g$, entonces $\int_A f \leq \int_A g$.*

v. *Si A tiene volumen y $|f| \leq M$, entonces $\left|\int_A f\right| \leq Mv(A)$.*

vi. **Teorema del valor medio para integrales** *Si $f: A \to \mathbb{R}$ es continua y A tiene volumen y es compacto y conexo, entonces existe $x_0 \in A$ tal que $\int_A f(x)\, dx = f(x_0)v(A)$. La cantidad $(1/v(A)) \int_A f$ se denomina **promedio** de f sobre A.*

vii. *Sea $f: A \cup B \to \mathbb{R}$. Si los conjuntos A y B son tales que $A \cap B$ tiene medida nula y $f|(A \cap B)$, $f|A$ y $f|B$ son todas integrables, entonces f es integrable en $A \cup B$ y $\int_{A \cup B} f = \int_A f + \int_B g$.*

Si $a < b < c$ en $\mathbb{R}$, la condición **vii** implica que

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx.$$

En el plano, si A y B son como se muestra en la figura 8.4-1, la integral sobre su unión es la suma de las integrales individuales, pues la intersección tiene medida nula (es un punto).

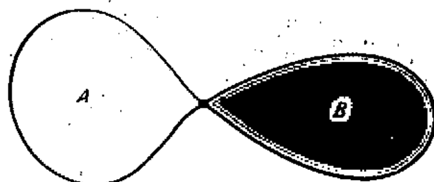


FIGURA 8.4-1 Dos conjuntos cuya intersección tiene medida nula

8.4.2 Ejemplo *Si A, B tienen volumen, muéstrese directamente (sin usar el teorema* **8.4.1***) que $A \cup B$ tiene volumen.*

Solución Debemos mostrar que $\partial(A \cup B)$ tiene medida nula (véase el corolario **8.3.2**). Pero

$$\partial(A \cup B) \subset \partial A \cup \partial B$$

(véase el ejercicio 15, capítulo 2), de modo que, como el miembro derecho tiene medida nula, también la tiene el miembro izquierdo. ♦

8.4.3 Ejemplo *Proporciónese una interpretación geométrica de la propiedad* **iii** *del teorema* **8.4.1** *para $f: [a, b] \to \mathbb{R}$.*

Solución $\int_a^b f(x)\,dx$ representa el área debajo de la gráfica de f, donde la parte por debajo del eje X se cuenta con signo negativo. La magnitud de esto es, claramente, menor (o igual) que el área por debajo de la gráfica de $|f|$; véase la figura 8.4-2. ♦

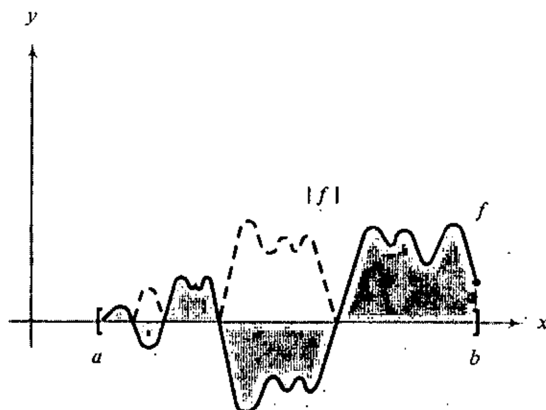


FIGURA 8.4-2 La integral de una función no es mayor que la de su valor absoluto

Ejercicios de §8.4

1. Si $A_1, A_2, \ldots$ tienen volumen y $A = A_1 \cup A_2 \cup \cdots$ está acotado, ¿debe A tener volumen?

2. Explíquese el significado geométrico de las propiedades **iv** y **v** en el teorema **8.4.1.**

3. Sean A, B conjuntos con volumen y $A \cap B$ de medida nula. Úsese **8.4.1** para mostrar que $v(A \cup B) = v(A) + v(B)$.

4. Sea $\int_a^b f(x)\,dx = -\int_a^b f(x)\,dx$ si $a > b$. Establézcase

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx$$

para todo a, b,c (no debe suponerse que $a < b < c$, como hemos hecho en el texto).

§8.5 Integrales impropias

Con frecuencia necesitamos integrar funciones no acotadas o integrar sobre regiones no acotadas. Las *integrales impropias* resultantes nos llevan al estudio de problemas de convergencia análogos a los de las series infinitas.

Por lo general, se definen las ***integrales impropias*** como

$$\int_0^{\infty} f(x)\,dx = \lim_{b\to\infty} \int_0^b f(x)\,dx,$$

o bien, si la función h no está acotada cerca de 0, como

$$\int_0^b h(x)\,dx = \lim_{\epsilon\to 0} \int_{\epsilon}^b h(x)\,dx,$$

y así sucesivamente. Las definiciones dadas en esta sección se ajustan a estos conceptos. Sin embargo, hay que tener cuidado: por lo general, *no* definimos

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{k\to\infty} \int_{-k}^{k} f(x)\,dx.$$

Si lo hiciéramos, considérese lo que ocurriría con $f(x) = x$: $\int_{-\infty}^{\infty} x\,dx$ sería igual a cero, mientras que $\int_0^{\infty} x\,dx$ y $\int_{-\infty}^{0} x\,dx$ no existirían; serían iguales a $+\infty$ y $-\infty$, respectivamente. Así, si queremos conservar la aditividad de las integrales, debemos proceder con más cuidado. Una forma de evitar esta "cancelación de infinitos" es separar f en sus partes positiva y negativa, como lo haremos en **8.5.4** más adelante.

En general, las integrales impropias son de dos tipos, dependiendo de si es la función o el dominio lo que no está acotado. En primer lugar, consideremos el caso de las regiones no acotadas. Para comenzar, supóngase que $f \geq 0$ está acotada y que $A \subset \mathbb{R}^n$ es un conjunto arbitrario, posiblemente no acotado. Extendemos f a todo el espacio de la manera usual, haciendo $f = 0$ fuera de A. Véase la figura 8.5-1.

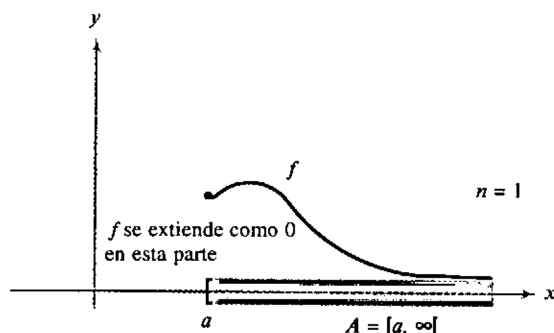


FIGURA 8.5-1 Una integral impropia en la recta real

8.5.1 Definición *Supóngase que $f(x) \geq 0$ para todo x en un conjunto $A \subset \mathbb{R}^n$ y que f está acotada y es integrable en cada "cubo" n-dimensional $[-a, a]^n = [-a, a] \times \cdots \times [-a, a]$. Si el límite $\lim_{a\to\infty} \int_{[-a,a]^n} f$ existe, lo llamamos $\int_A f$. Si existe y es finito, decimos que f es **integrable sobre** A.*

Aunque tomamos límites que se expanden simétricamente desde 0, hemos evitado el problema ya mencionado, pues solamente consideramos funciones no negativas. Más adelante estudiaremos las funciones con signo arbitrario.

8.5.2 Teorema *Supóngase que $f \geq 0$ y que está acotada y es integrable en cada cubo $[-a, a]^n$. Entonces f es integrable sii se cumple la siguiente condición: para cada sucesión B_k de conjuntos acotados con volumen tales que*

i. *$B_k \subset B_{k+1}$ para cada k y*

ii. *para cada cubo C tenemos $C \subset B_k$ para k suficientemente grande,*

$\lim_{k\to\infty} \int_{B_k} f$ existe. En este caso, $\lim_{k\to\infty} \int_{B_k} f = \int_A f$. Véase la figura 8.5-2.

Este teorema dice que podemos calcular $\int_A f$ independientemente de la forma de expansión hacia infinito.

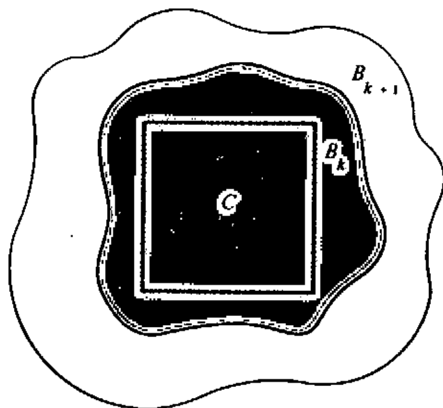


FIGURA 8.5-2 Una cadena creciente de conjuntos que absorben cada cubo

Obsérvese que tenemos un *criterio de comparación:* si $f \geq 0$ es integrable, g es integrable en todo cubo y $0 \leq g \leq f$, entonces g también es integrable, pues $\int_{[-a,a]^n} g$ es creciente con a y está acotada por la integral de f, por lo cual converge cuando $a \to \infty$.

A continuación analizaremos el caso de una función arbitraria $f \geq 0$ no acotada y definida en una región no acotada. La importancia de estas condiciones en el caso de $\mathbb{R}$ se verá en seguida. Reducimos el problema al caso ya analizado cortando la gráfica de f para obtener una función acotada.

8.5.3 Definición *Para cada número real positivo M, sea*

$$f_M(x) = \begin{cases} f(x), & si\, f(x) \leq M, \\ 0, & si\, f(x) > M \end{cases}$$

(véase la figura 8.5-3), de modo que f_M está acotada por M y $f_M \geq 0$ (obsérvese que $\int_A f_M$ crece si M crece, y que $0 \leq f_M \leq f$). Definimos

$$\int_A f = \lim_{M \to \infty} \int_A f_M$$

si este límite es finito, y en ese caso decimos que f es ***integrable***.

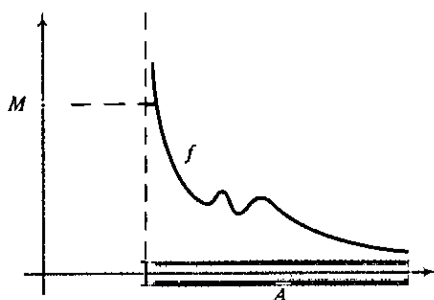


FIGURA 8.5-3 Truncamiento de una función no acotada

Como antes, si $f \geq 0$ es integrable y $0 \leq g \leq f$, entonces g también es integrable (éste es el ***criterio de comparación***).

8.5.4 Definición *Para una función general $f: A \to \mathbb{R}$ (es decir, una función f no necesariamente positiva), sean*

$$f^+(x) = \begin{cases} f(x), & si\, f(x) \geq 0, \\ 0, & si\, f(x) < 0, \end{cases}$$

y

$$f^-(x) = \begin{cases} -f(x), & si\, f(x) \leq 0, \\ 0, & si\, f(x) > 0 \end{cases}$$

(véase la figura 8.5-4). Si $\int_A f^+$ y $\int_A f^-$ existen, decimos que f es integrable y

$$\int_A f = \int_A f^+ - \int_A f^-.$$

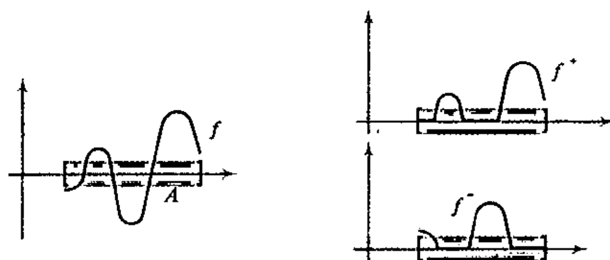


FIGURA 8.5-4 Las partes positiva y negativa de una función

Obsérvese que $|f| = f^+ + f^-$. Así, si f es integrable, también lo es $|f|$, y $\int_A |f| = \int_A f^+ + \int_A f^- \geq \left|\int_A f\right|$.

Si $|f|$ es integrable en cada cubo y si $\int |f|$ existe, entonces $\int f^+$ y $\int f^-$ deben ser finitas, si existen. Ambas son no negativas y cumplen $0 \leq f^+ \leq |f|$ y $0 \leq f^- \leq |f|$. Así, si $|f|$ es integrable y f es integrable en cada cubo, entonces f es integrable. Esta hipótesis sobre f es necesaria. Incluso en una región acotada es posible que $|f|$ sea integrable Riemann mientras que la propia f no lo sea (véase el ejercicio 33 al final del capítulo).

Aunque las integrales impropias son importantes en dimensiones mayores, el caso de la recta real merece una atención especial. En este caso, existe un método particularmente sencillo para calcular integrales.

8.5.5 Teorema

i. *Supóngase que* $f : [a, \infty[\to \mathbb{R}$ *es continua y que* $f \geq 0$. *Sea F una primitiva de f. Entonces f es integrable sii* $\lim_{x\to\infty} F(x)$ *existe. En este caso,*

$$\int_{[a,\infty[} f = \int_a^\infty f(x)\, dx = \left\{\lim_{x\to\infty} F(x)\right\} - F(a).$$

ii. *Supóngase que f: $]a, b] \to \mathbb{R}$ es continua y que $f \geq 0$. Entonces f es integrable sii*

$$\lim_{\varepsilon \to 0^+} \int_{a+\varepsilon}^{b} f(x)\, dx$$

existe. Este límite es igual a $\int_a^b f(x)\, dx$.

La generalización de esta afirmación a $\mathbb{R}^n$ aparece en el ejercicio 26 al final del capítulo. En vez de "$\int_A f$ existe", con frecuencia se dice "$\int_A f$ converge". Como en el caso de las series infinitas, es necesario disponer de un criterio de convergencia o divergencia para las integrales impropias. Algunos de estos criterios recuerdan los utilizados para series.

Uno de los criterios más útiles es el ***criterio de comparación***. Si $\int_a^\infty f(x)\, dx$ converge, $f \geq 0$, $0 \leq g \leq f$ y g es integrable en intervalos finitos, entonces $\int_a^\infty g(x)\, dx$ converge. La razón, como ya se estableció, es simplemente que $\int_a^b g(x)\, dx$ crece cuando $b \to \infty$ y está acotada superiormente por $\int_a^\infty f(x)\, dx$, por lo cual converge.

En nuestro estudio de las integrales impropias hemos utilizado lo que puede considerarse la extensión más natural a $\mathbb{R}^n$. Nuestro método es particularmente atractivo, pues la mayoría de las propiedades de la integral siguen siendo válidas. Sin embargo, en el caso particular de $\mathbb{R}^1 = \mathbb{R}$, también es útil considerar un tipo más débil de convergencia, llamada ***convergencia condicional***. En este caso, definimos

$$\int_a^\infty f(x)\, dx \text{ (condicional)} = \lim_{b \to \infty} \int_a^b f(x)\, dx$$

si el límite existe. Esto no es lo mismo que nuestra definición anterior (llamada ***convergencia absoluta***), pues en ella exigíamos que el límite existiese separadamente para f^+ y para f^-. Un ejemplo nos permitirá establecer la diferencia.

8.5.6 Ejemplo

Sea $f(x) = (\operatorname{sen} x)/x$. Muéstrese que $\int_1^\infty f(x)\, dx$ converge condicionalmente, pero no absolutamente.

Solución

Si f fuera integrable en $[1, \infty[$, entonces $|f|$ también lo sería. Sin embargo,

$$\int_1^\infty \frac{|\operatorname{sen} x|}{x}\, dx \geq \int_\pi^{n\pi} \frac{|\operatorname{sen} x|}{x}\, dx = \sum_{k=2}^{n} \int_{(k-1)\pi}^{k\pi} \frac{|\operatorname{sen} x|}{x}\, dx \geq \frac{2}{\pi} \sum_{k=2}^{n} \frac{1}{k}$$

pues en el intervalo $[(k-1)\pi, k\pi]$, $1/x \geq 1/k\pi$ y $\int_{(k-1)\pi}^{k\pi} |\operatorname{sen} x|\, dx = 2$. Pero $\sum_{k=2}^{n} 1/k \to \infty$ cuando $n \to \infty$, de modo que $\int_1^\infty |(\operatorname{sen} x)/x|\, dx = +\infty$.

Sin embargo, $\lim_{b\to\infty}\int_1^b((\operatorname{sen} x)/x)\,dx$ existe. Para verlo, obsérvese que una integración por partes da como resultado

$$\int_1^b \frac{\operatorname{sen} x}{x}\,dx = -\int_1^b \frac{d(\cos x)}{x} = -\frac{\cos b}{b} + \cos 1 - \int_1^b \frac{\cos x}{x^2}\,dx,$$

y $\int_1^\infty((\cos x)/x^2)\,dx$ existe, pues

$$\int_1^b \frac{|\cos x|}{x^2}\,dx \le \int_1^b \frac{1}{x^2}\,dx = 1 - \frac{1}{b},$$

que converge cuando $b\to\infty$. Así, $\int_1^\infty(\operatorname{sen} x)/x\,dx$ es condicionalmente convergente. ♦

Se pueden dar criterios refinados, como el criterio de Dirichlet para series, para demostrar la convergencia condicional cuando fallen los criterios para la convergencia absoluta. Véase el ejercicio 30 al final de este capítulo.

8.5.7 Ejemplo Daremos unas cuantas integrales impropias que resultan útiles junto con el criterio de comparación. Se puede demostrar su convergencia mediante la integración directa, una integración sucesiva por partes o mediante trucos similares.

a. $\int_1^\infty x^p\,dx \begin{cases}\text{converge si } p < -1\\ \text{diverge si } p \ge -1\end{cases}$

b. $\int_0^a x^p\,dx \begin{cases}\text{converge si } p > -1\\ \text{diverge si } p \le -1\end{cases}$

c. $\int_1^\infty e^{-x}x^p\,dx$ converge para todo p

d. $\int_0^a e^{1/x}x^p\,dx$ diverge para todo p

e. $\int_0^a \log x\,dx$ converge

f. $\int_1^\infty (1/\log x)\,dx$ diverge.

Por ejemplo, $\int_1^c x^p\,dx = x^{p+1}/(p+1)\big|_1^c$ si $p \neq -1$ y $\int_1^c(1/x)\,dx = \log x\big|_1^c$. Pero, $\log c \to \infty$ y $c^{p+1}\to\infty$ cuando $c\to\infty$ si $p+1>0$, es decir, si $p > -1$. Esto implica **a.** El apartado **b** es similar, y **e** y **f** se pueden demostrar de la misma forma. Bosquejamos la demostración de **c** en el ejercicio 1 al final de esta sección y **d** es análogo. ♦

8.5.8 Ejemplo *Muéstrese que* $\displaystyle\int_1^\infty \frac{1}{\sqrt{x^3+1}}\,dx$ *converge.*

Solución Esta integral es impropia cuando $x \to \infty$. Ahora, si $x \geq 0$,

$$\frac{1}{\sqrt{x^3+1}} \leq \frac{1}{\sqrt{x^3}} = x^{-3/2},$$

y $\int_1^\infty x^{-3/2}\,dx$ converge, por el apartado **a** de **8.5.7**. Por lo tanto, por comparación, esta integral también converge. ♦

Ejercicios de §8.5

1. Establézcase la fórmula **c** del ejemplo **8.5.7** como sigue. Demuéstrese que $e^{-x}x^{p+2} \to 0$ cuando $x \to \infty$ y después compárese la integral con $\int_1^\infty (1/x^2)\,dx$.

2. Sea $f: [a, \infty[\to \mathbb{R}$ integrable Riemann en intervalos acotados. Muéstrese que $\int_a^\infty f$ (convergencia condicional) existe sii para cada $\varepsilon > 0$ existe T tal que $t_1, t_2 \geq T$ impliquen

$$\left| \int_{t_1}^{t_2} f(x)\,dx \right| < \varepsilon$$

(éste es el ***criterio de Cauchy***).

3. Muéstrese que $\int_0^1 e^{-x}x^p\,dx$ converge si $-1 < p$.

4. Muéstrese que $\displaystyle\int_1^\infty \frac{\operatorname{sen} x}{x^2+1}\,dx$ es convergente.

5. ¿Para qué valores de α converge $\displaystyle\int_0^\infty \frac{x^\alpha}{1+x^\alpha}\,dx$?

§8.6 Algunos teoremas de convergencia

En el capítulo 5 vimos que la convergencia uniforme es suficiente para permitirnos intercambiar las operaciones de límite e integración. En esta sección perfeccionamos este resultado.

Un curso de este nivel no es el lugar adecuado para un estudio exhaustivo de los teoremas de convergencia, ya que encajan de forma más natural en los cursos avanzados de teoría de la medida e integración. Por lo tanto, nos restringiremos a un teorema ilustrativo: el teorema de la convergencia monótona, que necesitaremos en el capítulo 10 relativo a las series de Fourier. Para un análisis más completo de los teoremas de convergencia en la teoría de Riemann, véase W. A. J. Luxemburg, "Arzelá's Dominated Convergence Theorem for the Riemann Integral", *Am. Math. Monthly*, **78** (1971), págs. 970–979.

8.6.1 Teorema de la convergencia monótona de Lebesgue *Sea* $g_n : [0, 1] \to \mathbb{R}$ *una sucesión de funciones no negativas tales que cada integral impropia* $\int_0^1 g_n(x)\,dx$ *existe y es finita. Supóngase que* $0 \leq g_{n+1} \leq g_n$ *y que* $g_n(x) \to 0$ *para cada* $x \in [0, 1]$. *Entonces*

$$\lim_{n\to\infty} \int_0^1 g_n(x)\,dx = 0.$$

En este momento el lector debería regresar a los ejemplos de §**5.3** para ver que coinciden con este teorema. Ciertamente, si las funciones g_n no son decrecientes, el resultado no tiene por qué ser cierto, como lo muestran los ejemplos de esa sección.

8.6.2 Corolario *Sean* $f_n, f : [0, 1] \to \mathbb{R}$ *funciones integrables no negativas y supóngase que* $0 \leq f_n \leq f_{n+1} \leq f$ *y que* $\int_0^1 f(x)\,dx$ *existe como integral impropia. Además, supóngase que* $f_n(x) \to f(x)$ *para todo* $x \in [0, 1]$. *Entonces*

$$\lim_{n\to\infty} \int_0^1 f_n(x)\,dx = \int_0^1 f(x)\,dx.$$

Este resultado es consecuencia de la aplicación del teorema **8.6.1** a las funciones $g_n(x) = f(x) - f_n(x)$, que decrecen en forma monótona a cero. Los detalles se dejan al lector. Usamos el intervalo [0, 1] para fijar ideas, pero se podría usar cualquier otro intervalo, y los resultados también se podrían generalizar a $\mathbb{R}^n$.

8.6.3 Corolario *Si* $f : [a, b] \to \mathbb{R}$, $f \geq 0$ *y la integral impropia* $\int_a^b f^2 < \infty$ *existe, entonces* $\int_a^b (f - f_M)^2 \to 0$ *cuando* $M \to \infty$.

En este caso usamos el teorema **8.6.1** con $g_n = f - f_n$ (la función f_M, donde M se elige aquí como n, se definió en la sección anterior).

8.6.4 Ejemplo *Demuéstrese que* $\lim_{n\to\infty} \int_0^1 e^{-nx^2} x^p\,dx = 0$ *si* $p > -1$.

Solución Se aplica el teorema **8.6.1**. Las funciones $g_n(x) = e^{-nx^2} x^p$ satisfacen $g_n(x) \leq x^p$ y, por lo tanto, son integrables; además, las funciones g_n decrecen puntualmente a cero. ♦

8.6.5 Ejemplo *Sean g_n no negativas e integrables en* $[0, 1]$. *Sea* $g(x) = \sum_{n=1}^{\infty} g_n(x)$ *y supóngase que* $g(x)$ *es integrable. Demuéstrese que*

$$\int_0^1 g(x)\,dx = \sum_{n=1}^{\infty} \int_0^1 g_n(x)\,dx.$$

Solución Sea

$$f_n(x) = \sum_{k=1}^{n} g_k(x)dx, \quad \text{de modo que} \quad \int_0^1 f_n(x)\,dx = \sum_{k=1}^{n} \int_0^1 g_k(x)\,dx$$

Como las funciones $f_n(x)$ convergen de forma creciente a g, el corolario **8.6.2** muestra que sus integrales convergen a la integral de g. ♦

8.6.6 Ejemplo *Sea* $a_1 = 1$ *y* $1 \geq a_i \geq 0$, $i = 1, 2, \ldots,$ *una sucesión decreciente tal que* $a_i \to 0$. *Sea* $f_j \geq 0$ *una sucesión de números positivos y sea* $f(x) = f_j$ *en el intervalo* $]a_{i+1}, a_i]$, *como en la figura 8.6-1. Supóngase que f es integrable. Muéstrese que* $\int_0^1 f(x)\,dx = \sum_{n=1}^{\infty} f_j(a_j - a_{j-1})$.

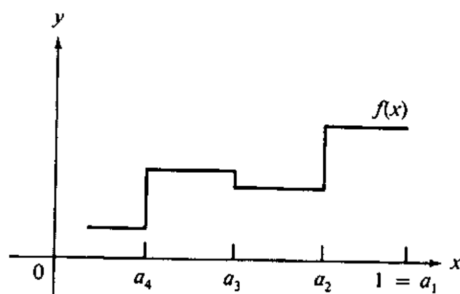


FIGURA 8.6-1 Una "función escalonada" con infinitos escalones

Solución Sea f_n la función que se anula si $x \leq a_n$ y es igual a $f(x)$ en caso contrario. Es claro que $\int_0^1 f_n(x)\,dx = \sum_{j=1}^{n} f_j(a_j - a_{j-1})$. Pero f_n es monótona creciente a f (excepto posiblemente en $x = 0$), de modo que de **8.6.2** obtenemos el resultado. ♦

Ejercicios de §8.6

1. Muéstrese que el teorema **8.6.1** se puede demostrar mediante los métodos del capítulo 5 si las funciones g_n son continuas.

2. Evalúese $\lim\limits_{n\to\infty} \int_0^1 \frac{e^x \operatorname{sen} nx}{n}\, dx$.

3. Evalúese $\lim\limits_{n\to\infty} \int_0^1 \frac{1 - e^{-nx}}{\sqrt{x}}\, dx$.

§8.7 Introducción a las distribuciones

Hacia 1930, en su famoso libro *Principios de mecánica cuántica*, Dirac enfatizó la utilidad de la función δ, que definió mediante las siguientes propiedades:

$$\delta(x) = \begin{cases} 0 & \text{si } x \neq 0, \\ \infty & \text{si } x = 0, \end{cases} \quad \text{y} \quad \int_{-\infty}^{\infty} \delta(x)\, dx = 1.$$

Uno puede imaginar aproximaciones a δ, donde $f_n \to \delta$ en cierto sentido (figura 8.7-1), pero difícilmente se puede visualizar δ directamente. Los físicos se dieron cuenta en seguida (como los ingenieros lo habían hecho independientemente) de la utilidad de tales ideas y procedieron a usar la función δ en sus cálculos. Por ejemplo, para describir la densidad de carga σ de una carga puntual con carga e, es conveniente escribir $\sigma = e\delta$. Uno concibe $e\delta$ como el límite de densidades de carga σ_n bien definidas, dispersas en áreas pequeñas que se concentran en un solo punto cuando $n \to \infty$.

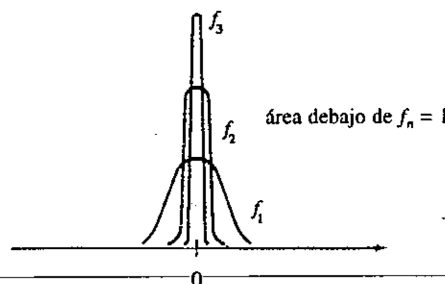


FIGURA 8.7-1 Una sucesión que aproxima la δ

Mientras los físicos y los ingenieros realizaban sus cálculos, los matemáticos se regodeaban "desde la barrera", señalando en ocasiones que todo el asunto relativo a esa

función δ carecía de sentido, pues no podía existir tal función. En realidad, la definición sí que carecía de sentido, como cualquiera podría ver. Para aumentar la diversión de los matemáticos, ¡Dirac procedió a derivar esta función δ!

Pero sucedió que los físicos habrían tenido una gran idea. En la actualidad, las distribuciones, de las que δ es un ejemplo, son indispensables en el análisis y las matemáticas aplicadas. Sin embargo, los matemáticos tardaron casi 20 años en establecer la teoría de distribuciones de una forma satisfactoria. Esto fue llevado a cabo por L. Schwartz y S.L. Sobolev cerca de 1948, aunque se habían encontrado indicios de la teoría en el trabajo de matemáticos anteriores. Echaremos sólo un breve vistazo a la teoría.

En los cálculos reales, δ casi siempre aparece debajo del signo de la integral en la forma

$$\int \delta(x)f(x)\,dx = f(0). \tag{1}$$

Podemos ver la idea detrás de la ecuación (1) a partir de la definición de Dirac, pues el hecho de que δ se anule fuera de $x = 0$ significa que solamente cuenta $f(0)$, de modo que deberíamos tener

$$\int \delta(x)f(x)\,dx = f(0)\int \delta(x)\,dx = f(0).$$

La ecuación (1) es la clave del modo de proceder, como sigue. Considérese el espacio $C_b(\mathbb{R},\ \mathbb{R})$ de funciones continuas acotadas en $\mathbb{R}$ (véase §5.5). Entonces, consideramos

$$\delta : C_b(\mathbb{R},\mathbb{R}) \to \mathbb{R}, \quad f \mapsto f(0).$$

Así, no consideramos δ como una función en $\mathbb{R}$ en absoluto, sino como una función en $C_b(\mathbb{R},\ \mathbb{R})$ que aplica f en $f(0)$. Esta operación está bien definida y es fácil ver que δ es continua (véase el ejercicio 40, capítulo 5).

Así, es posible evitar la dificultad de la definición de Dirac adoptando un punto de vista completamente nuevo; es decir, imaginar δ como una asignación de un número a cada función f. Esta asignación se denota mediante la expresión simbólica $\int \delta(x)f(x)\,dx$. Cualquier función continua g define también una operación de este tipo: aplica f en

$$\int_{-\infty}^{\infty} g(x)f(x)\,dx.$$

Así, las distribuciones (de las que son ejemplos las transformaciones lineales de C_b en $\mathbb{R}$) incluyen más que las funciones ordinarias.

¿Cómo derivar δ? Para esto, obsérvese que si g es derivable, entonces

$$\int_{-\infty}^{\infty} \frac{dg}{dx}f(x)\,dx = -\int_{-\infty}^{\infty} g(x)\frac{df}{dx}(x)\,dx,$$

siempre que f se anule para $|x|$ grande, como se puede ver mediante una integración

por partes. Así, dg/dx aplica f en el mismo número en que g aplica a $-df/dx$. Así, es lógico definir δ' como

$$\delta'(f) = \delta\left(-\frac{df}{dx}\right) = -\frac{df}{dx}(0).$$

Esto nos lleva a reemplazar C_b por las funciones f que son C^1 y se anulan para $|x|$ grande. También podemos reemplazar C_b por las funciones C^∞ que se anulan para $|x|$ suficientemente grande, lo que nos lleva a la siguiente definición.

8.7.1 Definición *Sea $\mathcal{D}$ el espacio de las funciones C^∞ que se anulan idénticamente fuera de cierto intervalo ($\mathcal{D}$ es conocido como el* ***espacio de las funciones prueba****). Una* ***distribución*** *T es una transformación lineal $T: \mathcal{D} \to \mathbb{R}$ (estrictamente, debe pedirse que T sea continua, en el sentido de que si $f_n \to f$ uniformemente en conjuntos acotados, todas las derivadas de f_n convergen uniformemente a las derivadas de f en conjuntos acotados y todas las funciones f_n se anulan fuera de un conjunto compacto fijo, entonces $T(f_n) \to T(f)$; sin embargo, la topología real de $\mathcal{D}$ es un tanto complicada). La* ***derivada*** *T' se define como $T' : \mathcal{D} \to \mathbb{R}$, $f \mapsto T(-df/dx)$.*

Si g es una función continua, se acostumbra a usar el mismo símbolo g para la distribución que aplica $f \mapsto \int_{-\infty}^{\infty} g(x) f(x)\, dx$.

Esta elegante formulación se debe a los fundadores de la teoría de las distribuciones. Todavía había que demostrar teoremas significativos acerca de las distribuciones, lo que supuso una importante revitalización de la teoría de las ecuaciones diferenciales en derivadas parciales. Los físicos quedaron encantados y todo el mundo contento (y así siguen, al menos por ahora, hasta que el escenario se repita con otros temas).

Ejercicios de §8.7

1. Muéstrese que $\delta''(f) = f''(0)$.

2. Sean T_n, T distribuciones. Digamos que $T_n \to T$ si $T_n(f) \to T(f)$ para todo $f \in \mathcal{D}$. Muéstrese que

$$\sqrt{\frac{n}{\pi}}\, e^{-nx^2} \to \delta.$$

3. Si $T_n \to T$ (véase el ejercicio 2), muéstrese que $T_n' \to T'$. Analícese y compárese esto con §**5.3.**

4. Encuéntrese una sucesión de funciones continuas g_n tales que $g_n \to \delta'$.

Demostraciones de los teoremas del capítulo 8

Demostraremos juntos el teorema de Darboux y la condición de Riemann.

8.1.2 Teorema de Darboux *Sea $A \subset \mathbb{R}^n$ acotado y contenido en algún rectángulo S. Sea $f: A \to \mathbb{R}$ acotada y extendámosla a S definiendo $f = 0$ fuera de A. Entonces f es integrable con integral I sii para todo $\varepsilon > 0$ existe $\delta > 0$ tal que si P es cualquier partición de S en rectángulos $S_1, \ldots, S_N$, con lados de longitud δ y si $x_1 \in S_1, \ldots, x_N \in S_N$, tenemos*

$$\left|\sum_{i=1}^{N} f(x_i)v(S_i) - I\right| < \varepsilon.$$

*Decimos que $\sum_{i=1}^{N} f(x_i)v(S_i)$ es una **suma de Riemann**.*

8.1.3 Condición de Riemann *f es integrable sii para todo $\varepsilon > 0$ existe una partición P_ε de S tal que $0 \leq U(f, P_\varepsilon) - L(f, P_\varepsilon) < \varepsilon$.*

Demostración Mostraremos que las condiciones "f integrable", "f satisface la condición de Riemann" y "f satisface la condición de Darboux" son equivalentes. Haremos esto en cuatro pasos.

Paso 1 *Si f es integrable, entonces f satisface la condición de Riemann.*

Demostración Dado $\varepsilon > 0$, existe una partición P_ε' tal que

$$U(f, P_\varepsilon') < I + \frac{\varepsilon}{2},$$

donde $I = \int_A f$. Podemos hacer esto, pues $I = \inf\{ U(f, P) \mid P \text{ es una partición}\}$. Si P es un refinamiento de P'_ε, entonces sabemos que

$$U(f,P) \leq U(f,P'_\varepsilon) < I + \frac{\varepsilon}{2}.$$

Análogamente elegimos P''_ε de modo que si P es un refinamiento de P''_ε, tenemos $L(f, P) > I - \varepsilon/2$. Sea $P_\varepsilon = P'_\varepsilon \cup P''_\varepsilon$. Si P es un refinamiento de P_ε, entonces

$$I - \frac{\varepsilon}{2} < L(f,P) \leq U(f,P) < I + \frac{\varepsilon}{2},$$

de modo que

$$0 \leq U(f,P) - L(f,P) < \varepsilon,$$

que es la condición de Riemann.

Paso 2 *Si f satisface la condición de Riemann, entonces f es integrable.*

Demostración Para cualquier $\varepsilon > 0$, existe P_ε tal

$$0 \leq U(f,P_\varepsilon) - L(f,P_\varepsilon) < \varepsilon.$$

Esto implica que $S = s$. De hecho, para cada P, tenemos

$$L(f,P) \leq s \leq S \leq U(f,P),$$

y entonces, si $U(f, P_\varepsilon) - L(f, P_\varepsilon) < \varepsilon$, también tenemos que $S - s < \varepsilon$ para cada $\varepsilon > 0$, y por lo tanto, $S = s$.

Paso 3 *Si f satisface la condición de Darboux, entonces f es integrable.*

Demostración Mostraremos que el valor I dado en la condición de Darboux será igual a $S = \inf\{ U(f, P) \mid P \text{ es una partición}\}$ y también igual a s. Para ello, dado $\varepsilon > 0$, produciremos una partición P tal que

$$|U(f,P) - I| < \varepsilon,$$

lo que mostrará que $S \leq I$. Análogamente, tendremos $I \leq s$ y entonces $I \leq s \leq S \leq I$ implicará $s = S = I$. Para hacerlo, elegimos $\delta > 0$ tal que si P es una partición con lados $< \delta$, entonces

$$\left|\sum f(x_i)v(S_i) - I\right| < \frac{\varepsilon}{2},$$

donde $S_1, \ldots, S_N$ son los rectángulos que conforman la partición P. Elíjase x_i tal que

$$|f(x_i) - \sup_{S_i}(f)| < \frac{\varepsilon}{v(S_i)2N}.$$

Entonces

$$|U(f, P) - I| \leq \left|U(f, P) - \sum f(x_i)v(S_i)\right| + \left|\sum f(x_i)v(S_i) - I\right|.$$

Ahora,

$$\left|U(f, P) - \sum f(x_i)v(S_i)\right| < \sum \frac{\varepsilon v(S_i)}{v(S_i)2N} = \frac{\varepsilon}{2},$$

de modo que $|U(f, P) - I| < \varepsilon$, como se pedía. El caso de las sumas inferiores es análogo.

Paso 4 *Si f es integrable, entonces f satisface la condición de Darboux.*

Demostración Supóngase que f es integrable, con integral I. Mostraremos en dos pasos que para cada $\varepsilon > 0$ existe $\delta > 0$ tal que si P es cualquier partición en rectángulos $S_1, \ldots, S_N$ de lados $< \delta$, y si $x_1 \in S_1, \ldots, x_N \in S_N$, tenemos

$$\left|\sum_{i=1}^{N} f(x_i)v(S_i) - I\right| < \varepsilon.$$

Paso 4A Sea P una partición del rectángulo $B \subset \mathbb{R}^n$. Dado $\varepsilon > 0$, mostraremos que *existe $\delta > 0$ tal que para cada partición P' en subrectángulos con lados menores que δ, la suma de los volúmenes de los subrectángulos de P' que no están totalmente contenidos en algún rectángulo de P es menor que ε.*

Para ello, examinaremos los casos $n = 1$ y $n > 1$ por separado. En primer lugar, supóngase que trabajamos con el intervalo $[a, b]$; supóngase que la partición P consta de N puntos. Afirmamos que la δ necesaria es simplemente ε/N. De hecho, la longitud de los intervalos de P' que no están contenidos en un intervalo de P es $N \times \delta$ = (número máximo de intervalos no totalmente contenidos en un intervalo de P) $\times$ (longitud máxima de cada uno de tales intervalos de P') $= \varepsilon$. En el caso general, la partición P consta de rectángulos $V_1, \ldots, V_M$. Denotamos T al "área" total de las caras situadas entre cualesquiera dos rectángulos. Sea $\delta = \varepsilon/T$ y P' cualquier partición de B en subrectángulos con lados menores que δ. Para cualquier rectángulo $S \in P'$ tal que S no esté contenido

en algún V_i, S interseca dos rectángulos adyacentes. Se puede ver que $v(S) \le \delta A$, donde A es el área total de las caras entre dos subrectángulos contenidos en S (véase la figura 8.P-1). Así, $\sum_{S\in P'} v(S) < \delta T = \varepsilon$.

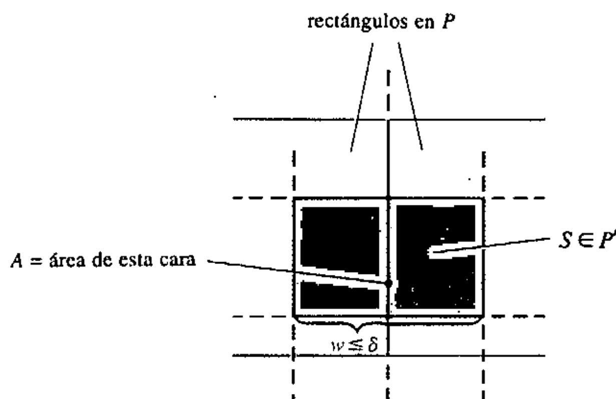


FIGURA 8.P-1 Demostración de que $v(S) = wA \le \delta A$

Paso 4B Como f es acotada, existe $M > 0$ tal que $|f(x)| < M$ para todo $x \in S$. Existen particiones P_1 y P_2 de S tales que $I - L(f, P_1) < \varepsilon/2$ y $U(f, P_2) - I < \varepsilon/2$. Elíjase una partición P que refine P_1 y P_2. Entonces $U(f, P) - I < \varepsilon/2$ e $I - L(f, P) < \varepsilon/2$. Por el paso **4A**, existe $\delta > 0$ tal que para toda partición de P en rectángulos con lados $< \delta$, la suma de los volúmenes de los subrectángulos no contenidos en algún subrectángulo de P es menor que $\varepsilon/2M$. Sea $S_1, \ldots, S_N$ una partición en subrectángulos con lados menores que δ; sean $S_1, \ldots, S_K$ subrectángulos contenidos en algún subrectángulo de P, y sean $S_{K+1}, \ldots, S_N$ los subrectángulos restantes. Si $x_1 \in S_1, \ldots, x_N \in S_N$, entonces

$$\begin{aligned}\sum_{i=1}^{N} f(x_i)v(S_i) &= \sum_{i=1}^{K} f(x_i)v(S_i) + \sum_{i=K+1}^{N} f(x_i)v(S_i)\\ &\le U(f, P) + M \cdot \frac{\varepsilon}{2M}\\ &= U(f, P) + \frac{\varepsilon}{2} < I + \varepsilon.\end{aligned}$$

Análogamente,

$$\sum_{i=1}^{N} f(x_i)v(S_i) \ge L(f, P) - \frac{\varepsilon}{2} > I - \varepsilon.$$

En consecuencia,

$$\left|\sum_{i=1}^{N} f(x_i)v(S_i) - I\right| < \varepsilon. \quad \blacksquare$$

En algunas demostraciones posteriores, será conveniente tener a la mano el siguiente detalle técnico: *en la definición de medida nula se pueden usar rectángulos abiertos o cerrados.*

Demostración Sea $A \subset \mathbb{R}^n$. En primer lugar, supóngase que, dado $\varepsilon > 0$, existen rectángulos abiertos $V_1, V_2, \ldots$ que recubren A con un volumen total $< \varepsilon$. Sea $B_i = \mathrm{cl}(V_i)$. Entonces $B_1, B_2, \ldots$ son rectángulos cerrados que recubren A con el mismo volumen total $< \varepsilon$.

Recíprocamente, dado $\varepsilon > 0$, supóngase que tenemos un recubrimiento mediante rectángulos cerrados $B_1, B_2, \ldots$ con volumen total $< \varepsilon/2^n$. Sea V_i el rectángulo abierto que contiene a B_i, con el doble de lado. Entonces $v(V_i) = 2^n v(B_i)$, de modo que

$$\sum_{i=1}^{\infty} v(V_i) = 2^n \sum_{i=1}^{\infty} v(B_i) < \varepsilon.$$

Este mismo argumento también sirve para el contenido nulo. Véase el ejercicio 11 al final del capítulo. ■

8.2.4 Teorema *Supóngase que los conjuntos $A_1, A_2, \ldots$ tienen medida nula en $\mathbb{R}^n$. Entonces $A_1 \cup A_2 \cup \cdots$ tiene medida nula en $\mathbb{R}^n$.*

Demostración Como todos los A_i tienen medida nula, existe un recubrimiento de A_i con rectángulos $B_{i1}, B_{i2}, \ldots$ tales que $\sum_{j=1}^{\infty} v(B_{ij}) < \varepsilon/2^i$. Como la colección $B_{i1}, B_{i2}, \ldots$ recubre los A_i, la colección numerable de B_{ij} recubre $A_1 \cup A_2 \cup \cdots$. Pero

$$\sum_{i,j=1}^{\infty} v(B_{ij}) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} v(B_{ij}) < \sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} = \varepsilon.$$

Como ε es arbitrario, $A_1 \cup A_2 \cup \cdots$ tiene medida nula. ■

Nota. El hecho de que podamos sumar los $v(B_{ij})$ primero con respecto a j y después con respecto a i es consecuencia de que los términos se pueden reordenar en una serie doble absolutamente convergente. Véase el ejercicio 51, capítulo 5.

8.3.1 Teorema de Lebesgue *Sean $A \subset \mathbb{R}^n$ un conjunto acotado y $f : A \to \mathbb{R}$ una función acotada. Extiéndase f a todo $\mathbb{R}^n$ definiéndola como cero en los puntos no con-*

tenidos en A. Entonces f es integrable (Riemann) sii los puntos donde la función f es discontinua forman un conjunto de medida nula.

Demostración Considérese algún rectángulo B que contenga a A. Entonces debemos mostrar que la función f es integrable en A sii el conjunto de discontinuidades de la función g, que es igual a f en A y cero fuera, tiene medida nula.

Para la demostración, es útil tener una medida de lo "mala" que puede ser una discontinuidad. Para esto, definimos la ***oscilación de una función h en x_0***, que se escribe $O(h, x_0)$, como

$$O(h, x_0) = \inf\{\sup\{|h(x_1) - h(x_2)| \mid x_1, x_2 \in U\} \mid U \text{ es una vecindad de } x_0\}.$$

Obsérvese que el ínfimo se toma sobre todas las vecindades U de x_0 y que $O(h, x_0) \geq 0$. Afirmamos que $O(h, x_0) = 0$ sii h es continua en x_0. Para verlo obsérvese que h es continua en x_0 sii para cada $\varepsilon > 0$ existe una vecindad U de x_0 tal que $\sup\{|h(x_0) - h(x_1)| \mid x_1 \in U\} < \varepsilon$, y que esto es equivalente a $O(h, x_0) = 0$.

Ahora estamos listos para continuar la demostración; por conveniencia, la separamos en dos pasos. Sea $g : B \to \mathbb{R}$ dada por $g(x) = f(x)$ si $x \in A$ y $g(x) = 0$ si $x \notin A$.

Paso 1 Supongamos que el conjunto de discontinuidades de g tiene medida nula. Así, si $D_\varepsilon = \{x \mid O(g, x) \geq \varepsilon\}$, para $\varepsilon > 0$ y $D = \{\text{discontinuidades de } g\}$, entonces $D_\varepsilon \subset D$. Si y es un punto de acumulación de D_ε, cada vecindad de y contiene un punto de D_ε. Entonces, cada vecindad U de y es una vecindad de un punto de D_ε, y por la construcción de D_ε, $\sup\{|f(x_1) - f(x_2)| \mid x_1, x_2 \in U\} \geq \varepsilon$. Esto implica $O(f, y) \geq \varepsilon$, por lo que $y \in D_\varepsilon$. Esto demuestra que D_ε es un conjunto cerrado. Puesto que $D_\varepsilon \subset B$, D_ε está acotado y, por lo tanto, es compacto. Ahora, D_ε es de medida nula, pues $D_\varepsilon \subset D$ y entonces, por definición, existe una colección $B_1, B_2, \ldots$ de rectángulos (abiertos) que recubren D_ε tales que $\sum_{i=1}^{\infty} v(B_i) < \varepsilon$. Sabemos que un número finito de B_i recubren D_ε, pues D_ε es compacto. Supóngase que $B_1, \ldots, B_N$ recubren D_ε. Está claro que $\sum_{i=1}^{N} v(B_i) < \varepsilon$.

Elíjase ahora una partición de B. Refinando esta partición podemos conseguir que cada rectángulo de ella o bien sea disjunto de D_ε o bien esté contenido en uno de los rectángulos $B_1, B_2, \ldots, B_N$ que recubren D_ε. Así, los rectángulos de la partición pertenecen a dos colecciones (no necesariamente disjuntas): $C_1 = \{\text{los rectángulos contenidos en uno de los } B_k\}$ y $C_2 = \{\text{los rectángulos que no intersecan a } D_\varepsilon\}$. Ahora usamos la compacidad para subdividir los rectángulos de C_2 y obtener otro refinamiento de nuestra partición. Para cada rectángulo S que no interseque a D_ε, la oscilación de g en cada punto del rectángulo es menor que ε. Por lo tanto, podemos hallar una vecindad U_x de cada punto x del rectángulo de modo que $M_{U_x}(g) - m_{U_x}(g) < \varepsilon$, donde $M_{U_x}(g) = \sup\{g(y) \mid y \in U_x\}$ y $m_{U_x}(g) = \inf\{g(y) \mid y \in U_x\}$. Como S es compacto, una colección finita de

los conjuntos abiertos U_x recubre S. Elíjase una partición refinada de S de modo que cada rectángulo de la partición esté contenido en U_{x_i} para algún U_{x_i} en la colección finita que recubre S. Si hacemos esto para cada S en C_2, obtenemos una partición P tal que

$$\begin{aligned} U(g,P) - L(g,P) &\leq \sum_{S\in C_1}(M_S(g) - m_S(g))v(S) + \sum_{S\in C_2}(M_S(g) - m_S(g))v(S) \\ &\leq \varepsilon v(B) + \sum_{S\in C_1} 2Mv(S), \text{ donde } |f(x)| < M \text{ en } A \\ &\leq \varepsilon v(B) + 2M\varepsilon, \qquad \text{pues } \sum_{S\in C_1} v(S) < \sum_{i=1}^{N} v(B_i) < \varepsilon. \end{aligned}$$

Como ε es arbitrario, la condición de Riemann muestra que g, y por lo tanto f, es integrable.

Paso 2 Supóngase que g es integrable. El conjunto de discontinuidades de g es el conjunto de puntos de oscilación mayor que cero. Por lo tanto, {discontinuidades de g} = $D_1 \cup D_{1/2} \cup D_{1/3} \cup \cdots$, donde $D_{1/n} = \{x \in B \mid O(g,x) \geq 1/n\}$. Por el teorema **8.1.2**, existe una partición de B tal que $U(g,P) - L(g,P) = \sum_{S\in P}(M_s(g) - m_s(g))v(S) < \varepsilon$. Por otro lado, $D_{1/n} = \{x \in D_{1/n} \mid x$ está en la frontera de algún $S\} \cup \{x \in D_{1/n} \mid x \in$ interior (S) para algún $S\} = S_1 \cup S_2$. El primero de estos conjuntos, S_1, tiene medida nula, pues podemos recubrir la frontera de un rectángulo con rectángulos arbitrariamente delgados. Sea C la colección de rectángulos de la partición que tienen un elemento de $D_{1/n}$ en su interior. Entonces, si $S \in C$,

$$M_S(g) - m_S(g) \geq \frac{1}{n}$$

y

$$\frac{1}{n}\sum_{S\in C'} v(S) \leq \sum_{S\in C}(M_S(g) - m_S(g))v(S) \leq \sum_{S\in P}(M_S(g) - m_S(g))v(S) < \varepsilon.$$

Por lo tanto, C es una colección de rectángulos que recubre S_2 y $\sum_{S\in C} v(S) < n\varepsilon$. Podemos encontrar una colección C' de rectángulos que recubre S_1 de modo que $\sum_{S\in C'} v(S) < \varepsilon$. Entonces $C \cup C'$ recubre $D_{1/n}$ y $\sum_{S\in C\cup C'} v(S) < (n+1)\varepsilon$. Como ε es arbitrario, $D_{1/n}$ tiene medida nula. Por último, {discontinuidades de g} $= D_1 \cup D_{1/2} \cup D_{1/3} \cup \cdots$ tiene medida nula, por el teorema **8.2.4.** ■

8.3.2 Corolario *Un conjunto acotado $A \subset \mathbb{R}^n$ tiene volumen sii la frontera de A tiene medida nula.*

Demostración Por el teorema **8.3.1**, basta demostrar que el conjunto de discontinuidades de 1_A, donde

$$1_A(x) = \begin{cases} 0, & \text{si } x \notin A \\ 1, & \text{si } x \in A, \end{cases}$$

es la frontera de A. Pero si $x \in \partial A$, entonces cualquier vecindad de x interseca a A y a $\mathbb{R}^n \setminus A$. Por lo tanto, existen puntos y en la vecindad tales que $1_A(x) - 1_A(y) = 1$. Así, 1_A no es continua en x. Si $x \notin \partial A$, entonces existe una vecindad de x que está totalmente contenida en A o en $\mathbb{R}^n \setminus A$. En cualquier caso, 1_k es constante en esa vecindad, de modo que 1_A es continua en x. ■

8.3.3 Corolario *Sea $A \subset \mathbb{R}^k$ acotado y con volumen. Una función acotada $f: A \to \mathbb{R}$ con una cantidad finita o numerable de puntos de discontinuidad es integrable.*

Demostración Las discontinuidades de la función extendida g, que es igual a f en A y cero en los puntos fuera de A, son sencillamente las discontinuidades de f, tal vez junto con algunas discontinuidades de g en la frontera de A, por la misma razón que en la demostración de **8.3.2.** Pero ∂A tiene medida nula, por el corolario **8.3.2**. Por lo tanto, basta mostrar que un conjunto numerable tiene medida nula. Pero esto es consecuencia del teorema **8.2.4** y del hecho de que un punto tiene medida nula. ■

8.3.4 Teorema

i. *Sea $A \subset \mathbb{R}^n$ un conjunto acotado, de medida nula y $f: A \to \mathbb{R}$ cualquier función integrable (acotada). Entonces $\int_A f(x)\, dx = 0$.*

ii. *Si $f: A \to \mathbb{R}$ es integrable y $f(x) \geq 0$ para todo x y $\int_A f(x)\, dx = 0$, entonces el conjunto $\{x \in A \mid f(x) \neq 0\}$ tiene medida nula.*

Demostración

i. Afirmamos que un conjunto de medida nula no puede contener un rectángulo no trivial, es decir, un rectángulo $[a_1, b_1] \times \cdots \times [a_n, b_n]$ tal que $a_i < b_i$ para cada i. La razón de esto es que un subconjunto de un conjunto de medida nula debe tener medida nula y un rectángulo no trivial no tiene medida nula. Sea S un rectángulo que contiene a A y extendamos f a S haciéndola igual a 0 en $S \setminus A$; sea P cualquier partición de S en subrectángulos $S_1, \ldots, S_N$ y sea M tal que $f(x) \leq M$ para todo $x \in A$.

Entonces

$$L(f,P)=\sum_{i=1}^{N} m_{S_i}(f)v(S_i) \le M\sum_{i=1}^{N} m_{S_i}(1_A)v(S_i).$$

Supóngase que $m_{S_i}(1_A) \neq 0$ para algún i, tal que S_i es un rectángulo (no trivial). Esto significa que $S_i \subset A$, lo que contradice las observaciones al principio de la demostración. Así, para cualquier S_i no trivial, $m_{S_i}(1_A) = 0$, mientras que si S_i es trivial, S_i, $v(S_i) = 0$. Por lo tanto, o $\sum_{i=1}^{N} m_{S_i}(1_A)v(S_i) = 0$, o $L(f, P) \le 0$. Ahora, $\sup_{x\in S_i} f(x) = -\inf_{x\in S_i}(-f(x))$, de modo que

$$U(f,P)=\sum_{S_i\in P}\sup_{x\in S_i} f(x)\, v(S_i) = -\sum_{S_i\in P}\inf_{x\in S_i}(-f(x))\, v(S_i) = -L(-f,P),$$

y de nuevo, por el mismo argumento, $L(-f, P) \le 0$. Por lo tanto, $-L(-f, P) = U(f, P) \ge 0$. Como P era arbitraria, $U(f, Q) \ge 0 \ge L(f, Q)$ para cualquier partición Q de S, por lo que

$$\overline{\int_A} f \ge 0 \ge \underline{\int_A} f,$$

y como f es integrable,

$$\overline{\int_A} f = \underline{\int_A} f = \int_A f = 0.$$

ii. Considérese el conjunto $A_m = \{x \in A \mid f(x) > 1/m\}$; primero mostraremos que A_m tiene contenido nulo. Dado $\varepsilon > 0$, sea S un rectángulo que contiene a A, extendamos f a S haciéndola igual a 0 en $S\backslash A$ y sea P una partición del rectángulo S tal que $U(f, P) < \varepsilon/m$. Dicha partición existe por el hecho de que $\int_A f = 0$. Si $S_1, \ldots, S_K$ son los subrectángulos de la partición P que tienen intersección no vacía con A_m y $M_{S_i}(f)$ es el supremo de f en S_i,

$$\sum_{i=1}^{K} v(S_i) \le \sum_{i=1}^{K} mM_{S_i}(f)v(S_i) < \varepsilon,$$

ya que $mM_{S_i}(f) > 1$. Por lo tanto, $S_1, \ldots, S_K$ es un recubrimiento por rectángulos cerrados del conjunto A_m tal que $\sum_{i=1}^{K} v(S_i) < \varepsilon$. Por lo tanto, A_k tiene contenido nulo. Como A_m tiene contenido nulo, también tiene medida nula.

Por último, obsérvese que

$$\{x \in A \mid f(x) \neq 0\} = \bigcup_{m=1}^{\infty} A_m.$$

Así, el teorema **8.2.4** implica que este conjunto tiene medida nula. ■

8.4.1 Teorema *Sean A, B subconjuntos acotados de $\mathbb{R}^n$, $c \in \mathbb{R}$ y f, $g : A \to \mathbb{R}$ funciones integrables. Entonces*

i. *$f + g$ es integrable y $\int_A (f + g) = \int_A f + \int_A g$.*

ii. *cf es integrable y $\int_A (cf) = c \int_A f$.*

iii. *$|f|$ es integrable y $|\int_A f| \leq \int_A |f|$.*

iv. *Si $f \leq g$, entonces $\int_A f \leq \int_A g$.*

v. *Si A tiene volumen y $|f| \leq M$, entonces $|\int_A f| \leq Mv(A)$.*

vi. **Teorema del valor medio para integrales** *Si $f : A \to \mathbb{R}$ es continua y A tiene volumen y es compacto y conexo, entonces existe $x_0 \in A$ tal que $\int_A f(x)\, dx = f(x_0)v(A)$. La cantidad*

$$\frac{1}{v(A)} \int_A f$$

*se denomina **promedio** de f sobre A.*

vii. *Sea $f : A \cup B \to \mathbb{R}$. Si los conjuntos A y B son tales que $A \cap B$ tiene medida nula y $f|(A \cap B)$ $f|A$ y $f|B$ son todas integrables, entonces f es integrable en $A \cup B$ y $\int_{A \cup B} f = \int_A f + \int_B g$.*

Nota. Si f es integrable en A y B, entonces es integrable en $A \cap B$. De hecho, si A y B tienen volumen, también $A \cap B$, pues $\partial(A \cap B) \subset \partial A \cup \partial B$.

Demostración

i. Sea S un rectángulo que contiene a A; extendemos f y g a S haciéndolas iguales a cero en $S \backslash A$. Sea $\varepsilon > 0$ dado. Por el teorema **8.1.2**, existe $\delta_1 > 0$ tal que si P_1 es cualquier partición de S en subrectángulos $S_1, \ldots, S_N$ con lados menores que δ_1 y si $x_1 \in S_1, \ldots, x_N \in S_N$, entonces

$$\left| \sum_{i=1}^{M} g(x_i)v(R_i) - \int_A g \right| < \frac{\varepsilon}{2}.$$

Análogamente, existe $\delta_2 > 0$ tal que si P_2 es cualquier partición de S en subrectángulos $R_1, \ldots, R_M$ con lados menores que δ_2 y si $x_1 \in R_1, \ldots, x_M \in R_M$, entonces

$$\left| \sum_{i=1}^{N} f(x_i)v(S_i) - \int_A f \right| < \frac{\varepsilon}{2}.$$

Si $\delta = \text{mín}(\delta_1, \delta_2)$, P es cualquier partición de S en subrectángulos $T_1, \ldots, T_K$ con lados menores que δ y $x_1 \in T_1, \ldots, x_K \in T_K$, por lo que

$$\left| \sum_{i=1}^{K} (g(x_i) + f(x_i)) v(T_i) - \int_A f - \int_A g \right| < \varepsilon.$$

Por el teorema **8.1.2**, podemos concluir que $f + g$ es integrable y que $\int_A (f+g) = \int_A f + \int_A g$.

ii. Sea $\varepsilon > 0$ dado. Sea S un rectángulo que contiene a A, con f extendida a S como en **i**. Sea $\delta > 0$ tal que si P es cualquier partición de S en subrectángulos $S_1, \ldots, S_N$ con lados menores que δ y si $x_1 \in S_1, \ldots, x_N \in S_N$, entonces

$$\left| \sum_{i=1}^{N} f(x_i) v(S_i) - \int_A f \right| < \frac{\varepsilon}{|c|}.$$

Esto implica que

$$\left| \sum_{i=1}^{N} c f(x) v(S_i) - c \int_A f \right| < \varepsilon.$$

Por lo tanto, cf es integrable y $\int_A (cf) = c \int_A f$.

iii. Demostraremos esta parte como un corolario de **iv**.

iv. Sea P cualquier partición del rectángulo S que contiene a A. Como $f \leq g$, vemos que $U(f, P) \leq U(g, P)$. Así,

$$\begin{aligned} \int_A f &= \inf\{ U(f,P) \mid P \text{ es cualquier partición} \} \\ &\leq \inf\{ U(g,P) \mid P \text{ es cualquier partición} \} = \int_A g. \end{aligned}$$

iii. Utilizamos el hecho de que si f es continua en un punto x de su dominio, entonces $|f|$ es continua en ese punto, pues $|f|$ es la composición de $y \mapsto |y|$ con f. En consecuencia, si f es integrable en A, el teorema **8.3.1** implica que $|f|$ es integrable en A. Ahora, $-|f| \leq f \leq |f|$ y por **iv**, $-\int_A |f| \leq \int_A f \leq \int_A |f|$, por lo que $\left| \int_A f \right| \leq \int_A |f|$.

v. Si P es cualquier partición del rectángulo S en subrectángulos $S_1, \ldots, S_N$, entonces

$$\int_A |f| \leq U(|f|, P) = \sum_{i=1}^{N} M_{S_i}(|f|) v(S_i) \leq M \sum_{i=1}^{N} M_{S_i}(1_A) v(S_i) = MU(1_A, P).$$

Esto implica que $\int_A |f| \leq Mv(A)$, y entonces

$$\left|\int_A f\right| \leq \int_A |f| \leq Mv(A).$$

vi. Sean $m = \inf\{f(x) \mid x \in A\}$ y $M = \sup\{f(x) \mid x \in A\}$. Por hipótesis, m y M son valores que se alcanzan, pues A es compacto. Sea $\lambda = (\int_A f)/v(A)$. (Si $v(A) = 0$, el teorema es consecuencia del **8.3.4.**) Por **v**, $m \leq \lambda \leq M$, y por el teorema de valores intermedios, existe $x_0 \in A$ tal que $f(x_0) = \lambda$, lo que demuestra la afirmación.

Nota. Un razonamiento un poco más cuidadoso muestra que la compacidad de A no es necesaria; véase el ejercicio 19.

vii. Sean $f_1 = f \cdot 1_A$, $f_2 = f \cdot 1_B$ y $f_3 = f \cdot 1_{A\cap B}$, de modo que esto represente las extensiones de $f|A$, $f|B$ y $f|A\cap B$ necesarias para la definición de integrabilidad. Por hipótesis, f_1, f_2 y f_3 son integrables y, por ejemplo, $\int_{A\cup B} f \cdot 1_A = \int_A f$, por definición. Ahora bien, $f = f_1 + f_2 - f_3$, por lo que **i** implica

$$\int_{A\cup B} f = \int_{A\cup B} f_1 + \int_{A\cup B} f_2 - \int_{A\cup B} f_3 = \int_A f + \int_B f - \int_{A\cap B} f.$$

Como $A \cap B$ tiene medida nula, el teorema **8.3.4** implica que $\int_{A\cup B} f = 0$. ■

8.5.2 Teorema *Supóngase que $f \geq 0$ y que está acotada y es integrable en cada cubo $[-a, a]^n$. Entonces f es integrable sii se cumple la siguiente condición: para cada sucesión B_k de conjuntos acotados con volumen tales que*

i. *$B_k \subset B_{k+1}$ para cada k y*

ii. *para cada cubo C tenemos $C \subset B_k$ para k suficientemente grande,*

$\lim_{k\to\infty} \int_{B_k} f$ *existe. En este caso,* $\lim_{k\to\infty} \int_{B_k} f = \int_A f$.

Demostración Supóngase que f es integrable. Si $[-a, a]^n \subset B_k \subset [-b, b]^n$ entonces

$$\int_{[-a,a]^n} f \leq \int_{B_k} f \leq \int_{[-b,b]^n} f,$$

pues $f \cdot 1_{[-a,a]^n} \leq f \cdot 1_{B_k} \leq f \cdot 1_{[-b,b]^n}$. Como $\lim_{a\to\infty} \int_{[-a,a]^n} f$ existe, también existe $\lim_{k\to\infty} \int_{B_k} f$, y el límite es el mismo. Por **ii**, para k suficientemente grande, tenemos $[-a, a]^n \subset B_k$ para cualquier a dado.

Para el recíproco, $\int_{B_k} f$ es una sucesión creciente en k, usando **i**, y tiene un límite; digamos que $\int_{B_k} f \to C$ cuando $k \to \infty$. Así, como $\int_{B_k} f \le C$ para todo k y como, para cada a, $[-a, a]^n \subset B_k$ para algún k, $\int_{[-a,a]^n} f \le C$ para todo a. Por lo tanto, cuando $a \to \infty$, ésta es una función de a creciente y acotada superiormente, por lo cual converge (¿por qué?). ■

8.5.5 Teorema

i. *Supóngase que* $f: [a, \infty[\to \mathbb{R}$ *es continua y que* $f \ge 0$. *Sea* F *una primitiva de* f. *Entonces* f *es integrable sii* $\lim_{x\to\infty} F(x)$ *existe. En este caso,*

$$\int_{[a,\infty[} f = \int_a^\infty f(x)\,dx = \left\{\lim_{x\to\infty} F(x)\right\} - F(a).$$

ii. *Supóngase que* $f:]a, b] \to \mathbb{R}$ *es continua y que* $f \ge 0$. *Entonces* f *es integrable sii*

$$\lim_{\varepsilon\to 0^+} \int_{a+\varepsilon}^b f(x)\,dx$$

existe. Este límite es igual a $\int_a^b f(x)\,dx$.

Demostración

i. $\int_{[-b,b]^n} f \cdot 1_{[a,b]} = \int_a^b f$ para $b > a$, y $\int_a^b f = F(b) - F(a)$. Por lo tanto, $\lim_{b\to\infty} \int_{-b}^b f$ existe sii $\lim_{b\to\infty} F(b)$ existe, y $\int_a^\infty f = (\lim_{b\to\infty} F(b)) - F(a)$.

ii. La segunda parte es un poco más enrevesada. Para $\varepsilon > 0$, definimos f^ε como 0 en $[a, a+\varepsilon]$ y f en $]a+\varepsilon, b]$. Primero damos dos pasos preliminares.

Paso 1 *Si* $M = \sup\{f(x) \mid a+\varepsilon \le x \le b$, *entonces, recordando que* $f_M(x) = f(x)$ *si* $f(x) \le M$ *y cero en caso contrario, tenemos*

$$\int_a^b f^\varepsilon = \int_{a+\varepsilon}^b f \le \int_a^b f_M,$$

pues $f \le f_M$ *en* $[a, b]$. Obsérvese que f_M podría ser distinta de cero en $[a, a+\varepsilon]$; en la figura 8.P-2 no lo es.

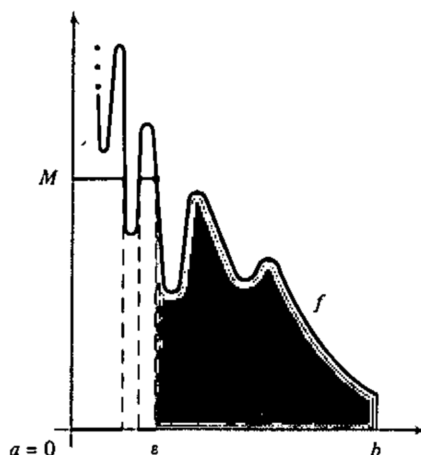


FIGURA 8.P-2 Corte y truncamiento de una función no acotada

Paso 2 *Para ε y M dados,*

$$\int_a^b f_M - \int_a^b f^\varepsilon \leq M\varepsilon.$$

Esto se debe a que $f_M \leq M$ en $[a, a+\varepsilon]$, y para $x \in [a+\varepsilon, b]$, $f_M(x) \leq f(x)$, de modo que

$$\int_a^b f_M - \int_a^b f^\varepsilon \leq \left(\int_a^{a+\varepsilon} f_M + \int_{a+\varepsilon}^b f\right) - \int_{a+\varepsilon}^b f = \int_a^{a+\varepsilon} f_M \leq \varepsilon M.$$

Para demostrar el teorema, primero supóngase que $\int f_M \to I$ cuando $M \to \infty$. Obsérvese que como $f \geq 0$, $\int f_M$ crece cuando M crece. Debemos mostrar que $\int f^\varepsilon$ también converge a I cuando $\varepsilon \to 0$. Está claro que crece hacia algo menor o igual que I, por el paso **1.** Dado $\delta > 0$, elíjase M de modo que $I - \int f_M < \delta/2$. Si $\varepsilon = \delta/2M$, entonces, por el paso **2**, $\int f_M - \int f^\varepsilon < \delta/2$. En consecuencia, $I - \int f^\varepsilon < \delta$. Así, $\lim_{\varepsilon \to 0} \int f^\varepsilon = I$.

El recíproco sigue un procedimiento análogo, usando de nuevo los pasos **1** y **2** para mostrar que si $\int f^\varepsilon \to I$, entonces $\int f_M = I$. ■

Como preparación para el teorema de la convergencia monótona, demostramos el siguiente lema.

Lema *Supóngase que f es integrable Riemann, $f: [0,1] \to \mathbb{R}$, $|f| \leq M$ y que $\int_0^1 f \geq \alpha > 0$. Entonces $E = \{x \in [0,1] \mid f(x) \geq \alpha/2\}$ contiene una unión finita de intervalos de longitud total $l \geq \alpha/4M$.*

Demostración Sea P una partición de $[0, 1]$ tal que $0 \leq \int_0^1 f - L(f, P) \leq \alpha/4$. Entonces $L(f, P) \geq 3\alpha/4$. Ahora mostraremos que los intervalos $R \in P$ tales que $R \subset E$ satisfacen la conclusión; sea l su longitud total. Tenemos que

$$\frac{3\alpha}{4} \leq L(f,P) = \sum_{R\in P, R\subset E} \inf_{x\in R} f(x)v(R) + \sum_{R\in P, R\not\subset E} \inf_{x\in R} f(x)v(R)$$

$$\leq Ml + \frac{\alpha}{2}(1-l) \leq Ml + \frac{\alpha}{2}.$$

En consecuencia, $l \geq \alpha/4M$, como se afirmaba. ■

8.6.1 Teorema de la convergencia monótona de Lebesgue *Sea $g_n : [0, 1] \to \mathbb{R}$ una sucesión de funciones no negativas tales que cada integral impropia $\int_0^1 g_n(x)\, dx$ existe y es finita. Supóngase que $0 \leq g_{n+1} \leq g_n$ y que $g_n(x) \to 0$ para cada $x \in [0, 1]$. Entonces*

$$\lim_{n\to\infty} \int_0^1 g_n(x)\,dx = 0.$$

Demostración Tenemos que $0 \leq \int_0^1 g_{n+1} \leq \int_0^1 g_n$; es decir, las integrales forman una sucesión acotada decreciente, de modo que $\lambda = \lim_{n\to\infty} \int_0^1 g_n$ existe y $\lambda \geq 0$. Queremos mostrar que $\lambda = 0$, de modo que suponemos que $\lambda > 0$. Obsérvese que $\int_0^1 g_n \geq \lambda$ para todo n. Considérese $E_n = \{x \in [0, 1] \mid g_n(x) \geq 2\lambda/5\}$ y obsérvese que $E_{n+1} \subset E_n$. Queremos aplicar el lema, pero g_n podría no estar acotada. Sin embargo, $g_n \leq g_1$; como $\int_0^1 g_1$ existe como integral impropia, $\int_0^1 g_{1M} \to \int_0^1 g_1$ cuando $M \to \infty$. En este caso,

$$g_{nM}(x) = \begin{cases} g_n(x) & \text{para } g_n(x) < M \\ M & \text{para } g_n(x) \geq M. \end{cases}$$

Considérese $M > 2\lambda/5$ y tal que $0 \leq \int_0^1 (g_1 - g_{1M}) \leq \lambda/5$ de modo que para todo n,

$$0 \leq \int_0^1 (g_n - g_{nM}) \leq \int_0^1 (g_1 - g_{1M}) \leq \lambda/5.$$

Por lo tanto, $\int_0^1 g_{nM} \geq 4\lambda/5 = \alpha$. Obsérvese que, como $M > 2\lambda/5$, E_n $\{x \in [0, 1] \mid g_{nM}(x) \geq 2\lambda/5\}$. En consecuencia, por el lema, E_n contiene una unión finita de intervalos de longitud total mayor o igual que $\lambda/5M$. Ahora definimos

$$D = \bigcup_{n=1}^{\infty} \{x \in [0, 1] \mid g_n \text{ no es continua en } x\};$$

entonces D tiene medida nula. Así, D está contenido en la unión U de una familia numerable de intervalos abiertos disjuntos de longitud total menor que $\lambda/5M$. Puede verse fácilmente que E_n no es un subconjunto de U. Obsérvese que si x_0 es un punto de acumulación de E_n, pero no está en E_n, entonces g_n debe ser discontinua en x_0, de modo que $x_0 \in D \subset U$. Es decir, $\text{cl}(E_n) \subset E_n \cup U$. Defínase $F_n = \text{cl}(E_n) \backslash U$. Por lo que acabamos de ver, $F_n \subset E_n$. Pero F_n es compacto y $F_{n+1} \subset F_n$, de modo que la propiedad de los conjuntos encajados implica que $\bigcap_{n=1}^{\infty} F_n \neq \varnothing$ y, por lo tanto, $\bigcap_{n=1}^{\infty} E_n \neq \varnothing$. Pero esto significa que para algún $x \in [0, 1]$, $g_n(x) \geq 2\lambda/5 > 0$ para todo n, lo que contradice la hipótesis $g_n(x) \to 0$. ■

8.6.3 Corolario *Si $f: [a, b] \to \mathbb{R}$, $f \geq 0$ y la integral impropia $\int_a f^2 < \infty$ existe, entonces $\int_a^b (f - f_M)^2 \to 0$ cuando $M \to \infty$.*

Demostración Aplicamos el teorema de la convergencia monótona a $g_n = (f - f_n)^2$. Como $(f - f_n + 1)^2 \leq (f - f_n)^2 \leq f^2$, las integrales $\int_a^b g_n$ existen y $g_{n+1} \leq g_n$. ■

Ejemplos resueltos del capítulo 8

Ejemplo 8.1 *Muéstrese que un conjunto acotado $A \subset \mathbb{R}^n$ tiene volumen nulo sii puede ser recubierto por un número finito de rectángulos de volumen total arbitrariamente pequeño.*

Solución Supóngase que A tiene volumen nulo y sea $\varepsilon > 0$ dado. Sea S un rectángulo cerrado que contiene a A y 1_A la función característica de A (es decir, es igual a 1 en A y a 0 en $S\backslash A$). Por definición de volumen nulo, existe una partición P de S en subrectángulos $S_1, \ldots, S_M$ tal que $U(1_A, P) < \varepsilon$. Sea P_0 la colección de aquellos subrectángulos S_i que intersecan A. Entonces $U(1_A, P) = \sum_{S \in P_0} v(S)$, de modo que P_0 es una colección de rectángulos que recubren A con un volumen total menor que ε.

Por otro lado, supóngase que dado $\varepsilon > 0$, A se puede recubrir con un número finito de rectángulos de volumen total menor que ε. Sean $V_1, \ldots, V_M$ tales rectángulos, S un rectángulo cerrado que contiene a A y P una partición de S en subrectángulos $S_1, \ldots, S_N$

tales que cada S_i está contenido en algún V_j o, a lo más, tiene frontera común con algunos V_i (definimos la partición usando todas las aristas de los V_j). Entonces $U(1_A, P) = \sum_{i=1}^{M} v(V_i) < \varepsilon$, lo que implica que $\inf\{U(1_A, P) \mid P \text{ es una partición}\} = 0$, de modo que $L(1_A, P') = 0$ para cualquier partición P', pues $0 \leq L(1_A, P) \leq U(1_A, P)$ para todo P. En consecuencia, A tiene volumen y dicho volumen es nulo. ♦

Ejemplo 8.2 *Sea f_k una sucesión de funciones integrables (Riemann) y acotadas definidas en $R = [a, b] \times [c, d]$. Supóngase que $f_k \to f$ uniformemente. Demuéstrese que f es integrable en R y que*

$$\iint_R f_k(x, y)\, dx\, dy = \iint_R f(x, y)\, dx\, dy$$

Solución (Compárese con el teorema **5.3.1**) Obsérvese primero que f está acotada. De hecho, si f_N es tal que $|f_N(x, y) - f(x, y)| < 1$ para todo $(x, y) \in R$, entonces, usando la desigualdad triangular,

$$|f(x, y)| \leq |f_N(x, y)| + 1.$$

Por lo tanto, como f_N está acotada, también lo está f.

Según el teorema **8.1.2**, debemos encontrar un número I tal que para todo $\varepsilon > 0$ exista $\delta > 0$ tal que

$$\left| \sum_{p,q=1}^{n,m} f(x'_p, y'_q)(x_p - x_{p-1})(y_q - y_{q-1}) - I \right| < \varepsilon$$

para cualquier subdivisión $x_0, \ldots, x_n$ de $[a, b]$ tal que $|x_p - x_{p-1}| < \delta$ y $x_{p-1} \leq x'_p \leq x_p$ y cualquier subdivisión análoga $y_0, \ldots, y_m$ de $[c, d]$. Así, *esperamos que* $I = \lim_{k\to\infty} \iint_R f_k(x, y)\, dx\, dy$. Ahora, $\iint_R f_k(x, y)\, dx\, dy$ es una sucesión de Cauchy. Para verlo, obsérvese que

$$\left| \iint_R f_k(x, y)\, dx\, dy - \iint_R f_l(x, y)\, dx\, dy \right| < \varepsilon \quad \text{si } |f_k(x, y) - f_l(x, y)| < \frac{\varepsilon}{\text{área }(R)}.$$

Por lo tanto, la sucesión converge a un valor que llamamos I.

Dado $\varepsilon > 0$, elegimos N tal que $k \geq N$ implica que $\left|\iint_R f_k(x, y)\, dx\, dy - I\right| < \varepsilon/3$; elíjase N_1 tal que si $k \geq N_1$ implica $|f_k(x, y) - f(x, y)| < \varepsilon/3(b-a)(d-c)$ para todo $(x, y) \in R$, y elíjase $N_2 = \text{máx}(N, N_1)$. Como f_{N_2} es integrable, existe $\delta > 0$ tal que para $|x_p - x_{p-1}| < \delta$, $|y_q - y_{q-1}| < \delta$,

$$\left| \sum_{p,q=1}^{n,m} f_{N_2}(x'_p, y'_q)(x_p - x_{p-1})(y_q - y_{q-1}) - \iint_R f_{N_2}(x)\, dx \right| < \frac{\varepsilon}{3}.$$

Con esta elección, la desigualdad triangular da por resultado

$$\left|\sum_{p,q=1}^{n,m} f(x'_p, y'_q)(x_p - x_{p-1})(y_q - y_{q-1}) - I\right|$$

$$\leq \left|\sum_{p,q=1}^{n,m} f(x'_p, y'_q)(x_p - x_{p-1})(y_q - y_{q-1}) - \sum_{p,q=1}^{n,m} f_{N_2}(x'_p, y'_q)(x_p - x_{p-1})(y_q - y_{q-1})\right|$$

$$+ \left|\sum_{p,q=1}^{n,m} f_{N_2}(x'_p, y'_q)(x_p - x_{p-1})(y_q - y_{q-1}) - \iint_R f_{N_2}(x, y)\, dx\, dy\right|$$

$$+ \left|\iint_R f_{N_2}(x, y)\, dx\, dy - I\right| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

lo que demuestra la afirmación. ♦

Obsérvese que este procedimiento se ha llevado a cabo sobre rectángulos de $\mathbb{R}^2$, pero que las mismas ideas son válidas sobre subconjuntos de $\mathbb{R}^n$. El lector debe formular la proposición precisa (véanse los ejercicios 14 y 16 al final de este capítulo).

Ejemplo 8.3 *Determínese* $\displaystyle\int_{[0,\infty[} \frac{dx}{(1+x)^2}$.

Solución Estamos integrando una función no negativa definida en un conjunto no acotado, por lo que, por definición,

$$\int_{[0,\infty[} \frac{dx}{(1+x)^2} = \lim_{a\to\infty} \int_0^a \frac{dx}{(1+x)^2}.$$

Como

$$\frac{d(-1/(1+x))}{dx} = \frac{1}{(1+x)^2},$$

el teorema fundamental muestra que

$$\int_0^a \frac{dx}{(1+x)^2} = -\frac{1}{1+a} + \frac{1}{1+0} = \frac{a}{1+a}.$$

Por lo tanto,

$$\lim_{a\to\infty} \int_0^a \frac{dx}{(1+x)^2} = 1 = \int_{[0,\infty[} \frac{dx}{(1+x)^2}. \quad ♦$$

Ejercicios del capítulo 8

1. **a.** Sea $f: A \subset \mathbb{R}^n \to \mathbb{R}$, donde A está acotado y f está acotada y es integrable en A. Considérese otra función acotada integrable $g: A \to \mathbb{R}$ tal que $g(x) = f(x)$ excepto en un conjunto $S \subset A$ de medida nula. Supóngase que f y g son integrables en S y en $A \backslash S$ y demuéstrese que $\int_A g = \int_A f$.

b. Si $f: A \subset \mathbb{R}^n \to \mathbb{R}$ y $g: A \subset \mathbb{R}^n \to \mathbb{R}$ son funciones acotadas, integrables en el conjunto acotado A y tales que $\int_A |f-g| = 0$, demuéstrese que $f(x) = g(x)$ para todo $x \in A$, excepto posiblemente para un conjunto de medida nula.

2. Proporciónese una demostración directa del teorema **8.4.1iii** por medio del teorema de Darboux y la desigualdad triangular para los números reales.

3. Demuéstrese que una función creciente $f: [a, b] \to \mathbb{R}$ es integrable Riemann.

4. Muéstrese que $f(x) = x^{2n+1}$ para n entero ≥ 0 no es (absolutamente) integrable sobre $\mathbb{R}$, aunque $\lim_{a\to\infty} \int_{-a}^{a} f$ existe.

5. En $\mathbb{R}^3$, demuéstrese que cualquier subconjunto del plano XY tiene medida nula.

6. Si $f: A \subset \mathbb{R}^n \to \mathbb{R}$ y $g: A \subset \mathbb{R}^n \to \mathbb{R}$ son funciones acotadas e integrables en el conjunto acotado A tales que $f(x) < g(x)$ para todo $x \in A$ y $v(A) \neq 0$, muéstrese entonces que $\int_A f < \int_A g$.

7. Sea $f: [a, b] \to \mathbb{R}$ continua y derivable en $]a, b[$. Supóngase que $f(a) = 0$, $f(b) = -1$ y $\int_a^b f(x)\,dx = 0$. Demuéstrese que existe $c \in\,]a, b[$ tal que $f'(c) = 0$.

8. Proporciónese una demostración del teorema **8.4.1i, ii, iii, iv** y **vii** para integrales impropias (las partes omitidas del teorema **8.4.1** no tienen sentido si A tiene volumen infinito).

9. Calcúlense las siguientes integrales:

a. $\int_0^{2\pi} \operatorname{sen} x\, dx$;

b. $\int_a^b x^2(\operatorname{sen} x)\, dx$.

10. Si S es un rectángulo abierto o cerrado, muéstrese que las dos definiciones de volumen coinciden; es decir, demuéstrese que $\int_S 1 = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n)$, donde $S = [a_1, b_1] \times \cdots \times [a_n, b_n]$ o $S =]a_1, b_1[\times \cdots \times]a_n, b_n[$.

11. **a.** Si $A \subset A_1 \cup A_2 \cup \cdots \cup A_N$ y todos los conjuntos A_i tienen volumen, muéstrese que $v(A) \leq \sum_{i=1}^{N} v(A_i)$.

b. Si A es compacto, muéstrese que A tiene medida nula sii A tiene contenido nulo. Muéstrese que un conjunto acotado B tiene volumen sii ∂B tiene contenido nulo.

12. Demuéstrese que A tiene medida nula sii para todo $\varepsilon > 0$ existe un recubrimiento de A por conjuntos $V_1, V_2, \ldots$ con volumen, tales que $\sum_{i=1}^{\infty} v(V_i) < \varepsilon$.

13. Demuéstrese que una función acotada $f: S \to \mathbb{R}$ es integrable en el rectángulo S sii existe un número I tal que para cualquier $\varepsilon > 0$ existe una partición P_ε tal que para todo refinamiento P de P_ε y cualquier elección de $x_i \in S_i$ con $S_i \in P$, $\left|\sum_{S_i = P} f(x_i)v(S_i) - I\right| < \varepsilon$.

14. Generalícese el ejemplo resuelto **8.2** al caso de las funciones $f: A \subset \mathbb{R}^n \to \mathbb{R}$.

15. ¿Es $\int_0^\infty x^p\,dx$ convergente para todo número real p? En caso contrario, ¿para qué valores de p?

16. a. Supóngase que $f_k \to f$ uniformemente en $A \subset \mathbb{R}^n$. Sea A_k el conjunto de los puntos de discontinuidad de f_k (extendida). Muéstrese que las discontinuidades de f (extendida) están contenidas en $A_1 \cup A_2 \cup \cdots$.

b. Úsese **a** para dar otra demostración de la integrabilidad de la función límite del ejemplo **8.2**.

c. Encuéntrense funciones integrables $f_k : A \to \mathbb{R}$ tales que $f_k \to f$ puntualmente pero que f no sea integrable.

17. a. Sea P_n la división de $[a, b]$ en 2^n subintervalos de igual longitud y formemos $L(f, P_n)$ y $U(f, P_n)$ para $f: [a, b] \to \mathbb{R}$ una función acotada. Muéstrese que f es integrable sii $\lim_{n\to\infty} L(f, P_n) = \lim_{n\to\infty} U(f, P_n)$. ¿Por qué siempre existen dichos límites?

b. Generalícese **a** para $A \subset \mathbb{R}^n$.

18. a. Generalícese el teorema del valor medio para integrales (**8.4.1vi**) al caso en que A sea un conjunto acotado y conexo arbitrario.

b. Si $\varphi(x) \geq 0$ para x en un conjunto compacto y conexo $A \subset \mathbb{R}$; φ es continua y creciente en x, y f es positiva e integrable, entonces $f\varphi$ es integrable y $\int_A f\varphi = \varphi(c)\int_A f$ para algún punto $c \in A$ (segundo teorema del valor medio).

c. Muéstrese que **b** falla si A está acotado pero no es compacto.

19. Sea $f: B \to \mathbb{R}$ integrable y $f \geq 0$. Si $A \subset B$ y f es integrable en A, entonces $\int_A f \leq \int_B f$. ¿Es cierto esto si no suponemos que $f \geq 0$?

20. Supóngase que $f:]0, b] \to \mathbb{R}$ es continua, positiva e integrable en $]0, b]$, y que, cuando $x \to 0$ por la derecha, $f(x)$ crece monótonamente a $+\infty$. Demuéstrese que $\varepsilon f(\varepsilon) \to 0$ cuando $\varepsilon \to 0$.

21. Muéstrese que $\int_0^\infty x^{-p}$ sen $x\,dx$ converge si $p > 1$. Muéstrese que si $0 < p \leq 1$, entonces la convergencia es condicional.

22. La ***función gamma*** se define como la función dada por la integral impropia $\Gamma(p) = \int_1^\infty e^{-x} x^{p-1}\,dx$. Muéstrese que la integral converge si $p > 0$.

23. **a.** Si $\varphi : [a, b] \to \mathbb{R}^n$ es una función continua, muéstrese que el conjunto $S =$ gráfica de $\varphi = \{(x, \varphi(x)) \mid x \in [a, b]\} \subset \mathbb{R}^{n+1}$ tiene contenido y medida nulos.

b. Para $\varphi : \mathbb{R} \to \mathbb{R}^n$ continua, muéstrese que la gráfica de φ tiene medida nula.

c. Sea $f : [a, b] \to \mathbb{R}$ integrable. Muéstrese que la gráfica de φ tiene volumen nulo, considerando la diferencia entre las sumas superior e inferior de f.

d. Muéstrese que la elipse $x^2 + 3y^2 = 9$ en $\mathbb{R}^2$ tiene volumen nulo.

24. Proporciónese un ejemplo para mostrar que lo siguiente no es equivalente a la integrabilidad de f: para cualquier $\varepsilon > 0$ existe $\delta > 0$ tal que si P es cualquier partición en rectángulos $S_1, \ldots, S_N$ con lados $< \delta$, existe $x_1 \in S_1, \ldots, x_N \in S_N$ tal que

$$\left| \sum_{i=1}^{N} f(x_i) v(S_i) - I \right| < \varepsilon.$$

25. Sea $f : [0, \infty[\to \mathbb{R}$ una función no negativa, integrable y uniformemente continua. Muéstrese que

$$\lim_{x \to \infty} f(x) = 0.$$

26. Considérese un conjunto $A \subset \mathbb{R}^n$, donde A está acotado y tiene volumen. Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$, $f \geq 0$ tal que $f \geq 0$, donde f puede no estar acotada. Supóngase que C_i es una sucesión de conjuntos compactos con volumen tales que $C_i \subset A$ y C_i crecen hacia A; además, $v(C_i) \to v(A)$ (esto es cierto automáticamente, como veremos en el ejercicio 15 al final del capítulo 9). Muéstrese que f es integrable sii f es integrable en cada C_i, $\lim_{i\to\infty} \int_{C_i} f$ existe y

$$\int_A f = \lim_{i \to \infty} \int_{C_i} f.$$

27. Demuéstrese que si $f : A \subset \mathbb{R}^n \to \mathbb{R}$ es continua, A es abierto con volumen y $\int_B f = 0$, para cada $B \subset A$ con volumen, entonces $f = 0$.

28. Sea $f : [0, 1] \to \mathbb{R}$ integrable en $[0, 1]$ y continua en x_0. Muéstrese que la transformación $x \mapsto \int_0^x f(y)\,dx$ es derivable en x_0 con derivada igual a $f(x_0)$. Proporciónese

un ejemplo de una función f integrable y discontinua para la cual esta transformación no sea derivable. Para una función f acotada e integrable, demuéstrese que esta transformación siempre es continua y, de hecho, es Lipschitz.

29. Muéstrese que el conjunto de Cantor $C \subset [0, 1]$ tiene medida nula (véase el ejercicio 38 al final del capítulo 3).

30. Demuéstrense los siguientes análogos de los criterios de Weierstrass y Dirichlet para la convergencia uniforme, por medio del criterio de Cauchy:

a. Sea $f: [a, \infty[\times [c, d] \to \mathbb{R}$ y supóngase que existe una función positiva $M(x)$ definida para $x \in [a, \infty[$, tal que $|f(x, s)| \leq M(x)$ para todo $s \in [c, d]$, y tal que $\int_a^\infty M(x)\, dx < \infty$. Entonces $F(s) = \int_a^\infty f(x, s)\, dx$ converge uniformemente en s. Si $f(x, s)$ es continua en (x, s), demuéstrese que F es continua.

b. Sea $f: [a, \infty[\times [c, d] \to \mathbb{R}$ continua y supóngase que $\left|\int_a^r f(x, s)\, dx\right| \leq M$, con M constante para todo $r \geq a$, $s \in [c, d]$. Supóngase que $\varphi(x, s)$ es decreciente en x y que $\varphi(x, s) \to 0$ cuando $x \to \infty$ uniformemente en s. Demuéstrese que $F(s) = \int_a^\infty \varphi(x, s) f(x, s)\, dx$ converge uniformemente.

31. Sea $f: [a, b] \to \mathbb{R}$ derivable y supóngase que f' es integrable. Demuéstrese que $\int_a^b f'(x)\, dx = f(b) - f(a)$.

32. Supóngase que φ es una función derivable en $[a, b]$ y que f es continua en la imagen de φ. Muéstrese que

$$\frac{d}{dx}\left[\int_{\varphi(a)}^{\varphi(x)} f(t)\, dt\right] = f(\varphi(x))\varphi'(x).$$

33. Defínase una función f en el intervalo $[0, 1]$ haciendo $f(x) = 1$ si x es racional y $f(x) = -1$ si x es irracional. Muéstrese que $|f|$ es integrable en $[0, 1]$ pero que f no lo es.

34. Defínase una función f en $[0, 1]$ haciendo $f(x) = 1$ si x es irracional y $f(x) = 1/q$ si x es un número racional igual a p/q para p, q enteros sin factores comunes. ¿Es f integrable en $[0, 1]$? (Véase también el ejercicio 38.)

35. Sea $A_n = [(n+1) + (n+2) + \cdots + (n+n)]/n$. Demuéstrese que $\lim_{n\to\infty}(1/n)A_n = 3/2$ por medio de la integral de Riemann.

36. Demuéstrese que $\lim_{n\to\infty}(n!)^{1/n}/n = e^{-1}$, considerando las sumas de Riemann para $\int_0^1 \log x\, dx$ basadas en la partición $1/n < 2/n < \cdots < 1$.

37. **a.** ¿En qué condiciones $\int_a^b f(\varphi(t))\, \varphi'(t)\, dt = \int_{\varphi(a)}^{\varphi(b)} f(x)dx$?

b. Evalúese $\int \frac{dx}{(1-x)\sqrt{1-x^2}}$ usando $x = \cos t$.

38. Sea $f\colon [0, 1] \to \mathbb{R}$ dada por

$$f(x) = \begin{cases} 0 & \text{si } x \text{ es irracional} \\ 1/q & \text{si } x = p/q, \end{cases}$$

donde p, $q > 0$ no tienen factores comunes. Muéstrese que f es integrable y calcúlese $\int_0^1 f$ (véase también el ejercicio 34).

39. Demuéstrese que $\log 2 = \lim_{n\to\infty} \left[\frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n}\right]$.

40. Encuéntrese un subconjunto abierto de $\mathbb{R}$ contenido en $]0, 1[$ que no tenga volumen como sigue:

a. Revísese la construcción del conjunto de Cantor (véase el ejercicio 39, capítulo 3).

b. Modifíquese el conjunto de Cantor haciendo que C_k se obtenga de C_{k-1} eliminando el intervalo central de tamaño $1/2^k$ de cada intervalo de C_{k-1} y haciendo $C_0 = [0, 1]$. Sea $C = \bigcap_{k=1}^{\infty} C_k$.

c. Muéstrese que $v(C_k) = \prod_{i=1}^{k}(1 - 1/2^i) \geq \frac{1}{4}$.

d. Sea U el complementario de C. Determínese la frontera de U y, usando **c**, muéstrese que no puede tener medida nula.

Este ejercicio produce también un ejemplo de conjunto compacto C con interior vacío que no tiene volumen.

41. Sea $R([a, b]) = \{f : [a, b] \to \mathbb{R} \mid f \text{ es Riemann integrable}\}$. Sea

$$d(f, g) = \int_a^b |f(x) - g(x)|\, dx.$$

¿Es d una métrica en el espacio $R([a, b])$?

42. Encuéntrese un subconjunto A de $[0, 1]$ tal que $A = \mathrm{cl}(\mathrm{int} A)$ y aun así ∂A no tenga medida nula.

43. Es un hecho (véase J. Marsden y M. Hoffman, *Basic Complex Analysis*, 2ª ed., W.H. Freeman, Nueva York, 1987, pág. 26) que

$$\operatorname{sen}\frac{\pi}{n}\operatorname{sen}\frac{2\pi}{n}\cdots\operatorname{sen}\frac{(n-1)\pi}{n} = \frac{n}{2^{n-1}}.$$

Úsese esta identidad para evaluar la integral impropia $\int_0^{\pi} \log \operatorname{sen} x\, dx$.

44. Analícese la generalización del teorema **8.6.1** a $\mathbb{R}^n$.

45. Analícese la generalización del teorema **8.6.1** de $[0, 1]$ a $[0, \infty[$.

46. **a.** Supóngase que $U =]-1, 1\,[\times]\,-1, 1[\, \subset \mathbb{R}^2, f: U \to \mathbb{R}$, que $\partial f/\partial x$ y $\partial f/\partial y$ existen en cada punto de U y que están acotadas en U, donde (x, y) son las coordenadas usuales de $\mathbb{R}^2$. Muéstrese que f es continua en $(0, 0)$.

b. Muéstrese, mediante un ejemplo, que el hecho de que las derivadas parciales estén acotadas es necesario en la parte **a**; la simple existencia no es suficiente.

47. Para cada $\alpha > 0$, compárese $\int_0^N x^\alpha\, dx$ con $\sum_{n=1}^{N} n^\alpha$ y $\sum_{n=0}^{N-1} n^\alpha\, dx$ y con ello determínese

$$\lim_{N\to\infty} \sum_{n=1}^{N} \frac{n^\alpha}{N^{1+\alpha}}.$$

48. Para cualquier función $f(x)$ continua en los reales, defínase la sucesión $f_n(x) = n\int_x^{x+1/n} f(\xi)d\xi$ para $n = 1, 2, 3, \ldots$. Muéstrese que $df_n(x)/dx$ existe aunque $df(x)/dx$ no exista, que $f(x) = \lim_{n\to\infty} f_n(x)$ y que la convergencia al límite es uniforme cuando f es uniformemente continua.

49. Supóngase que $\{I_k\}$ es una colección de intervalos abiertos cuya unión recubre un intervalo compacto C en el eje real; muéstrese que existe un número positivo ε tal que cada subintervalo de C que no sea más ancho que ε esté totalmente contenido en al menos uno de los I_k.

50. Establézcanse los lemas, teoremas, etcétera, necesarios para justificar cada una de las siguientes afirmaciones:

a. $\lim_{n\to\infty} \sum_{k=1}^{\infty} 2^{-k}\text{sen}(k/n) = 0.$

b. Si $f(x)$ es una serie de potencias convergente en $]-1, 1[$, entonces ocurre lo mismo para $f'(x)$.

c. Sean $f(x) = \tan(\pi x/2)$ y $a_n = f^{(n)}(0)/n!$. Entonces $\sum_{n=0}^{\infty} a_n$ no es una serie convergente (no se intente calcular a_n).

d. Si $f_n(x)$ es derivable en $[a, b]$ y $f_n'(x)\,| < 10$ para todo n y si $x \in [a, b]$ y $f_n(x) \to 0$ en cada x, entonces $f_n(x) \to 0$ uniformemente.

e. $f(x) = \sum_{k=1}^{\infty} \dfrac{\cos(2^k x)}{3^k}$ tiene derivada continua.

f. $\left|e^{x-1} - x - x^2/2 - \cdots - x^{99}/99!\right| \le e^x x^{100}/100!$ para $x > 0$

g. $f(x) = \sum_{n=1}^{\infty} x^n/n^3$ es continua en el intervalo cerrado $[-1, 1]$.

h. $\lim_{x\to 0} \dfrac{e^{ax} - 1}{\text{sen}\, bx} = \dfrac{a}{b}.$

i. Si $\sum_{n=0}^{\infty} |a_n| < \infty$, entonces $\int_0^t (\sum_{n=0}^{\infty} a_n \cos nx)\, dx = \sum_{n=0}^{\infty} (a_n/n) \operatorname{sen} nt$.

j. Para algún entero n, $n > 10^{100}(\log n)^{1000}$.

51. ¿Cierto o falso? Una función definida en $[0, \infty[$ pero que no es infinitamente derivable no se puede expresar como la suma de una serie de Dirichlet

$$f(t) = \sum_{n=0}^{\infty} a_n e^{-nt}.$$

52. Sea Ω el conjunto de puntos $(x, y) \in \mathbb{R}^2$ que se pueden expresar en la forma $(\operatorname{sen}\theta + \operatorname{sen}\psi, \cos\theta - \cos\psi)$ para algún par (θ, ψ). Encuéntrese un punto interior de Ω (es decir, un punto que, junto con un disco en torno de él, pertenece a Ω).

53. Muéstrese que la serie

$$\sum_{k=1}^{\infty} \frac{1}{2^k - k \operatorname{sen}(kx)}$$

es uniformemente convergente en $\mathbb{R}$.

54. Demuéstrese el "teorema de la convergencia dominada para series". Si

i. para cada k, $a_{k,n} \to a_k$ cuando $n \to \infty$, a_k y $a_{k,n} \in \mathbb{R}^m$,

ii. para cada n y k, $\|a_{k,n}\| \le b_k$, $\|a_k\| \le b_k$ para algún $b_k \in \mathbb{R}$, y

iii. $\sum_{k=1}^{\infty} b_k$ es convergente (es decir, por el criterio de comparación, $\sum_{k=1}^{\infty} a_k$ y $\sum_{k=1}^{\infty} a_{k,n}$ son convergentes), entonces

$$\sum_{k=1}^{\infty} a_{k,n} \to \sum_{k=1}^{\infty} a_k \text{ cuando } n \to \infty.$$

Proporciónese un ejemplo para mostrar que la condición **iii** es necesaria, aun cuando se suponga que $\sum_k a_{k,n}$ y $\sum_k a_k$ son convergentes.

55. Encuéntrese una función $f(x)$ derivable tal que $f'(x)$ esté acotada pero no sea integrable en $[0, 1]$.

Capítulo 9
Teorema de Fubini y la fórmula del cambio de variables

§9.1 Introducción

Dos teoremas fundamentales de integración nos ayudan a evaluar las integrales múltiples. El primero se refiere a la evaluación de integrales múltiples por medio de una integración iterada; con este método podemos calcular el valor de una integral múltiple mediante varias integraciones simples.

9.1.1 Ejemplo *Sea A el cuadrado dado por* $0 \leq x \leq 1$ *y* $0 \leq x \leq 1$*; evalúese la inte-gral de* $(x + y)x$ *sobre A.*

Solución

$$\int_A (x + y)x\,dx\,dy = \int_{x=0}^{1} \left(\int_{y=0}^{1} [x^2 + yx]\,dy \right) dx$$
$$= \int_0^1 \left(x^2 + \frac{1}{2}x \right) dx = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}. \quad \blacklozenge$$

Si A no es un cuadrado, sino, por ejemplo, un triángulo, entonces extendemos la función a un cuadrado haciéndola cero fuera de A. Entonces, en el proceso de integración mostrado en el ejemplo, la integración con respecto a y se corta en algún punto que depende de x, como se indica en la figura 9.1-1.

La idea intuitiva detrás del método del ejemplo **9.1.1** es la siguiente: para una función dada $f(x, y)$, $0 \leq x \leq 1$, $0 \leq y \leq 1$, el número $\int_0^1 f(x, y)\,dy$ es el área debajo de la gráfica de f sobre la recta x = constante; al integrar esta área sobre x obtenemos el

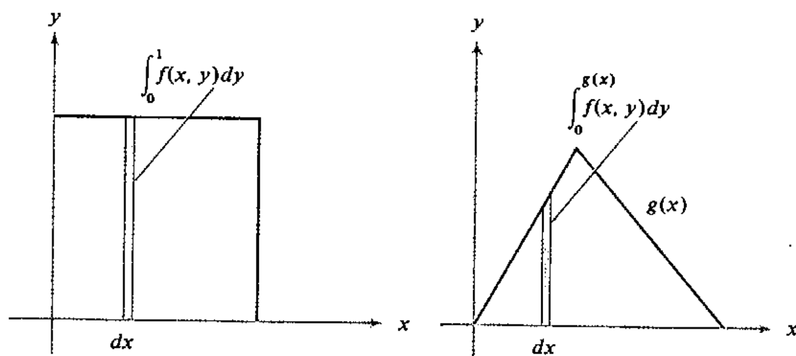


FIGURA 9.1-1 Integrales iteradas

volumen debajo de la gráfica. Esto sugiere que $\int_A f(x, y)\,dx\,dy = \int_0^1 (\int_0^1 f(x, y)\,dy)dx$. El resultado preciso se dará en el teorema de Fubini **9.2.1.** ♦

El segundo teorema básico del capítulo es la *fórmula del cambio de variables*, que se usa junto con el método del ejemplo **9.1.1** para evaluar ciertos tipos de integrales. El siguiente ejemplo es típico.

9.1.2 Ejemplo *Evalúese* $\int_A (x^2 + y^2)\,dx\,dy$, *donde A es el disco unidad dado por* $A = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$.

Solución Cambiamos de variables a las coordenadas polares $(r, \theta) : x = r\cos\theta$, $y = r\operatorname{sen}\theta$, para $r > 0$ y $0 < \theta < 2\pi$. Tenemos la "regla" $dx\,dy = r\,dr\,d\theta$. Como $x^2 + y^2 = r^2$, obtenemos

$$\int_A (x^2 + y^2)\,dx\,dy = \int_A r^2 r\,dr\,d\theta = \int_{r=0}^{1}\int_{\theta=0}^{2\pi} r^3\,d\theta\,dr = \int_{r=0}^{1} 2\pi r^3\,dr = \frac{\pi}{2}. \quad ♦$$

La justificación de la "regla" $dx\,dy = r\,dr\,d\theta$ está dada por la fórmula del cambio de variables (teorema **9.3.1**). El factor adicional r es justamente el jacobiano $\partial(x, y)/\partial(r, \theta)$. Sin embargo, es fácil "justificar" heurísticamente la regla, considerando dr y $d\theta$ como infinitesimales: dr representa un infinitesimal radial, mientras que $r d\theta$ representa una longitud de arco infinitesimal generadas por un cambio $d\theta$ en el ángulo a lo largo de un círculo de radio r. Así, $r\,dr\,d\theta$ es el elemento de área en un sector acotado por r, $r + dr$ y θ, $\theta + d\theta$ (véase la figura 9.1-2).

En una dimensión, la fórmula del cambio de variables es sencilla; establece que si f es continua en $[a, b]$ y tenemos una transformación y $\varphi : [\alpha, \beta] \to [a, b]$ tal que $\varphi(\alpha) = a$,

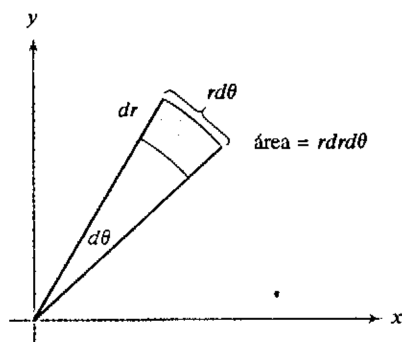


FIGURA 9.1-2 Un "elemento de área" en coordenadas polares

$\varphi(\beta) = b$ y φ' existe y es continua, entonces

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f(\varphi(u))\varphi'(u)\,du.$$

Para demostrarlo, encontramos primero una F tal que $F' = f$, lo que es posible por el teorema fundamental, y observamos que

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Ahora, sea $G(u) = F(\varphi(u))$. Por la regla de la cadena,

$$G'(u) = F'(\varphi(u))\varphi'(u) = f(\varphi(u))\varphi'(u).$$

y usando de nuevo el teorema fundamental,

$$\int_\alpha^\beta f(\varphi(u))\varphi'(u)\,du = G(\beta) - G(\alpha) = F(b) - F(a),$$

como se pedía (este teorema también es válido si f no es continua sino solamente integrable, como veremos en §9.3). Esta técnica se llama con frecuencia "integración por sustitución" y el estudiante ya conoce su eficacia:

9.1.3 Ejemplo *Intégrese* $\int (1 + x^2)^{10}\, x\, dx$.

Solución Sea $y = 1 + x^2$ y obsérvese que $y' = 2x$, de modo que

$$\int (1 + x^2)^{10} x\,dx = \frac{1}{2}\int y^{10} y'\,dx = \frac{1}{2}\int y^{10}\,dy = \frac{1}{22}y^{11} + C = \frac{1}{22}(1 + x^2)^{11} + C,$$

donde C es una constante. ♦

La generalización de este resultado a dimensiones mayores aparece en el enunciado de la fórmula del cambio de variables; aunque su demostración es bastante más sutil, es fácil de comprender y, junto con el teorema de Fubini, es uno de los métodos de cálculo más poderosos con los que contamos.

Ejercicios de §9.1

1. Evalúese $\int_0^1 x^2 e^{x^3}\, dx$.

2. Evalúese $\displaystyle\int_0^1 \frac{\operatorname{sen} x}{\cos^2 x}\, dx$.

3. Evalúese $\int_A (x + y^2)\, dx\, dy$, donde $A = [0, 1] \times [0, 1]$.

4. Evalúese $\int_A e^{-x^2-y^2}\, dx\, dy$, donde A es el disco unidad en $\mathbb{R}^2$.

5. Evalúese $\int_A dx\, dy$, donde A es el triángulo en $\mathbb{R}^2$ acotado por las rectas $x = 0$, $y = 0$, $x + y = 1$.

§9.2 Teorema de Fubini

Enunciaremos ahora el primero de nuestros dos teoremas básicos. Comenzaremos con el teorema de Fubini para el caso del plano $\mathbb{R}^2$.

9.2.1 Teorema

i. *Sea A el rectángulo descrito por $a \le x \le b$, $c \le y \le d$ y sea $f : A \to \mathbb{R}$ continua. Entonces*

$$\int_A f = \int_a^b \left(\int_c^d f(x,y)\, dy \right) dx = \int_c^d \left(\int_a^b f(x,y)\, dx \right) dy.$$

La expresión

$$\int_a^b \left(\int_c^d f(x,y)\, dy \right) dx$$

significa que la función

$$g(x) = \int_c^d f(x,y) dy dy$$

se integra de a a b.

ii. *En* **i**, *supóngase que f es integrable y que la función $f_x : [c, d] \to \mathbb{R}$ definida por $f_x(y) = f(x, y)$ es integrable para cada $x \in [a, b]$ fijo. Entonces*

$$\int_A f = \int_a^b \left(\int_c^d f(x,y)\, dy \right) dx.$$

De manera análoga, podemos suponer que

$$\int_a^b f(x,y)\, dx$$

existe para cada y, y obtener

$$\int_A f = \int_c^d \left(\int_a^b f(x,y)\, dx \right) dy.$$

Como es usual, podemos aplicar esto a una región no cuadrada A, extendiendo f como cero fuera de A y aplicando este teorema a un rectángulo que la contenga. Obsérvese que al hacer esta extensión, f puede desarrollar discontinuidades en la frontera de A.

No se puede eliminar la hipótesis de existencia de las integrales *iteradas* y reemplazarlas por la existencia de la integral *doble* de f en el presente contexto. Para obtener dicho resultado, el estudiante deberá esperar a cursos más avanzados de teoría de la medida. Sin embargo, lo habitual es que, en la práctica, el teorema **9.2.1** resulte adecuado. Como mencionamos en §**9.1**, el resultado es, intuitivamente, completamente razonable.

El siguiente corolario es una aplicación típica del teorema. Este corolario se puede usar con frecuencia si se descompone una región complicada en regiones más pequeñas a las cuales se aplica el corolario.

9.2.2 Corolario *Sean $\varphi, \psi : [a, b] \to \mathbb{R}$ transformaciones continuas tales que $\varphi(x) \leq \psi(x)$ para todo $x \in [a, b]$; sea*

$$A = \{(x,y) \mid a \leq x \leq b, \varphi(x) \leq y \leq \psi(x)\},$$

y sea $f : A \to \mathbb{R}$ continua. Entonces (véase la figura 9.2-1)

$$\int_A f = \int_a^b \left(\int_{\varphi(x)}^{\psi(x)} f(x,y)\, dy \right) dx.$$

Existe un teorema análogo en el que se intercambian los papeles de x e y. El corolario es una consecuencia del teorema si recordamos que f se extiende como cero fuera de A. Estos resultados se extienden a integrales múltiples como sigue:

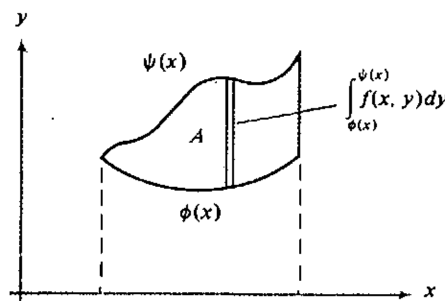


FIGURA 9.2-1 Teorema de Fubini

9.2.3 Teorema

i. Sean $A \subset \mathbb{R}^n$ *y* $A \subset \mathbb{R}^m$ *rectángulos y sea* $f: A \times B \subset \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ *una función continua. Para cada* $x \in A$, *defínase* $f_x : B \subset \mathbb{R}^m \to \mathbb{R}$ *como* $fx(y) = f(x, y)$. *Entonces*

$$\int_{A\times B} f = \int_A \left(\int_B f(x,y)\,dy \right) dx = \int_A \left(\int_B f_x(y)\,dy \right) dx.$$

ii. *Si f es integrable y f_x es integrable para cada x fijo, entonces, nuevamente,*

$$\int_{A\times B} f = \int_A \left(\int_B f(x,y)\,dy \right) dx.$$

Análogamente, si $\int_A f(x, y)\,dx$ existe para cada y, entonces

$$\int_{A\times B} f = \int_B \left(\int_A f(x,y)\,dx \right) dy.$$

En la práctica, este teorema se puede usar varias veces para reducir un problema a integrales iteradas de dimensión 1.

9.2.4 Ejemplo *Evalúese*

$$\int_A (x + y + z)^2\,dx\,dy\,dz$$

donde A es el volumen tridimensional representado en la figura 9.2-2.

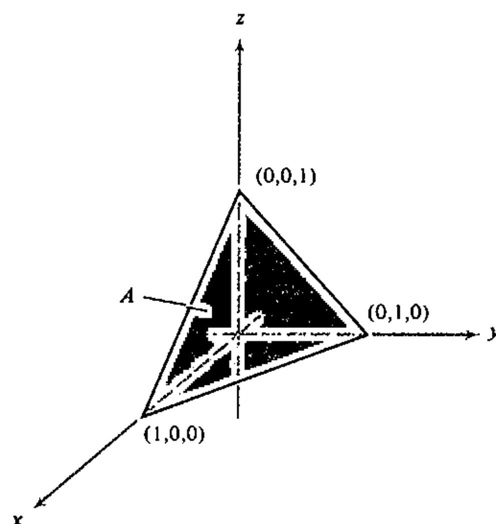


FIGURA 9.2-2 Una región tetraédrica en $\mathbb{R}^3$

Solución En este caso, A es el conjunto $\{(x, y, z) \in \mathbb{R}^3 \mid x \geq 0, y \geq 0, z \geq 0 \text{ y } x + y + z \leq 1\}$, por lo que consta de aquellos puntos (x, y, z) para los cuales $x \geq 0, y \geq 0, x + y \leq 1$ y $0 \leq z \leq 1 - (x + y)$. Sea $B = \{(x, y) \mid x \geq 0, y \geq 0 \text{ y } x + y \leq 1\}$. Entonces, por el teorema **9.2.3**, y recordando que f se anula fuera de A,

$$\int_A (x + y + z)^2 \, dx\, dy\, dz = \int_B \left(\int_0^{1-(x+y)} (x + y + z)^2 \, dz \right) dx\, dy.$$

Análogamente, B consta de los puntos (x, y) para los que $x \in [0, 1]$ e $y \in [0, 1 - x]$, de modo que

$$\int_A (x + y + z)^2 \, dx\, dy\, dz = \int_0^1 \int_0^{1-x} \int_0^{1-(x+y)} (x + y + z)^2 \, dz\, dy\, dx.$$

Usamos ahora el teorema fundamental para evaluar estas integrales. En primer lugar, obsérvese que

$$\frac{\partial}{\partial z} \left[\frac{1}{3}(x + y + z)^3 \right] = (x + y + z)^2,$$

de modo que al realizar la integración con respecto a z obtenemos

$$\int_A (x+y+z)^2\,dx\,dy\,dz$$
$$= \int_0^1 \int_0^{1-x} \left(\frac{(x+y+1-(x+y))^3}{3} - \frac{(x+y+0)^3}{3} \right) dy\,dx$$
$$= \int_0^1 \int_0^{1-x} \left(\frac{1}{3} - \frac{(x+y)^3}{3} \right) dy\,dx.$$

De nuevo, usando el teorema fundamental, la integración con respecto a y da como resultado

$$\int_A (x+y+z)^2\,dx\,dy\,dz = \int_0^1 \left(\frac{1-x}{3} - \frac{(x+(1-x))^4}{12} + \frac{x^4}{12} \right) dx$$
$$= \int_0^1 \left(\frac{1}{3} - \frac{x}{3} - \frac{1}{12} + \frac{x^4}{12} \right) dx$$
$$= \frac{1}{3} - \frac{1}{6} - \frac{1}{12} + \frac{1}{60} = \frac{20-10-5+1}{60} = \frac{1}{10}. \quad \blacklozenge$$

Para las integrales impropias, por lo general basta aplicar primero el teorema en una región acotada y después tomar los límites. Véase el ejemplo 1 al final de este capítulo.

9.2.5 Ejemplo *En la siguiente integral, cámbiese el orden de integración y evalúese:*

$$\int_0^1 \int_x^1 xy\,dy\,dx.$$

Solución La región en cuestión se muestra en la figura 9.2-3 (véase el corolario **9.2.2**). En el orden original, primero integramos con respecto a y. Para cada x entre 0 y 1, y va de x a 1, como en el apartado (a) de la figura. En el orden inverso, para cada y entre 0 y 1, x va de 0 a y, como en el apartado (b) de la figura. Obtenemos

$$\int_0^1 \int_0^y xy\,dx\,dy = \int_0^1 \frac{y^3}{2}\,dy = \frac{1}{8}. \quad \blacklozenge$$

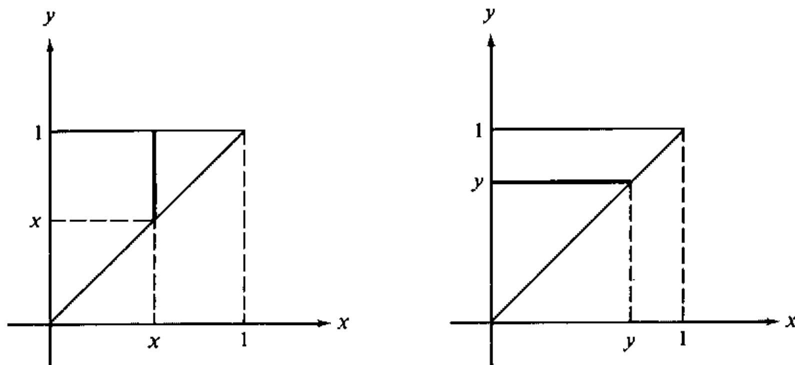


FIGURA 9.2-3 Una región triangular de integración en $\mathbb{R}^2$

Ejercicios de §9.2

1. Muéstrese que el volumen de la región A en el ejemplo **9.2.4** es 1/6.
2. Dibújese la región correspondiente a la integral $\int_0^1 \int_1^{e^x} (x + y)\, dy\, dx$ y evalúese.
3. Intercámbiese el orden de integración en el ejercicio 2 verificando que la respuesta no se altera.
4. Sea A la región en $\mathbb{R}^3$ acotada por los planos $x = 0$, $y = 0$, $z = 2$ y la superficie $z = x^2 + y^2$. Muéstrese que

$$\int_A x\, dx\, dy\, dz = 8\left(\frac{\sqrt{2}}{15}\right).$$

§9.3 Teorema del cambio de variables

Ahora veremos la fórmula del cambio de variables para integrales múltiples.

9.3.1 Teorema del cambio de variables *Sean $a \subset \mathbb{R}^n$ un conjunto abierto acotado con volumen y sea $g : A \to \mathbb{R}^n$ una transformación C^1 inyectiva. Supóngase que*

$Jg(x) \neq 0$ para todo $x \in A$ y que $|Jg(x)|$ y $1/|Jg(x)|$ están acotados en A. Sea $B = g(A)$ y supóngase que B tiene volumen (véase **§8.2**). *Si $f: B \to \mathbb{R}$ está acotada y es integrable, entonces $(f \circ g)|J(g)|$ es integrable en A y*

$$\int_B f = \int_A (f \circ g)|Jg|,$$

es decir,

$$\int_B f(y_1, \ldots, y_n)\, dy_1 \cdots dy_n = \int_A f(g(x_1, \ldots, x_n)) \frac{\partial(g_1, \ldots, g_n)}{\partial(x_1, \ldots, x_n)}\, dx_1 \cdots dx_n.$$

La demostración de este teorema necesita cierta manipulación sutil, pero podemos dar un argumento sencillo para su plausibilidad, como sigue.

Para comenzar, supóngase que aislamos un pequeño rectángulo S en A. Entonces, g es aproximadamente afín cerca de S, de modo que $g(S)$ es aproximadamente un paralelepípedo, como en la figura 9.3-1. Si g fuera afín, el volumen de $g(S)$ sería $|\det \tilde{g}|\, v(S)$, donde $\tilde{g}$ es la parte lineal de g. Sin embargo, $g(x_0) + \mathbf{D}g(x_0)$ aproxima bien g cerca de x_0 y es afín, de modo que $v(g(S)) \approx |Jg|\, v(S)$. Así,

$$f(g(x))|Jg(x)|dx \approx f(y)dy,$$

donde $y = g(x)$, por lo que al "sumar" estas cantidades infinitesimales obtenemos el resultado.

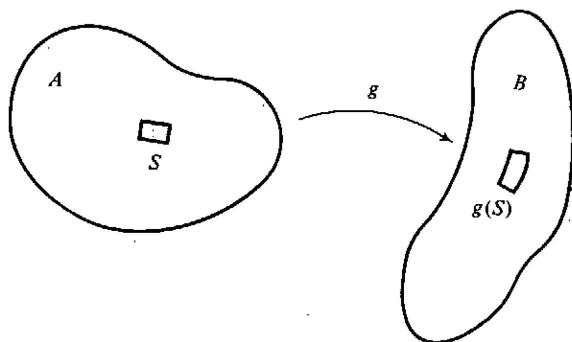


FIGURA 9.3-1 Cambio de coordenadas

9.3.2 Ejemplo *¿Cuál es el volumen del paralelepípedo generado por los vectores $(1, 1, 1)$, $(2, 3, 1)$, $(0, 1, 1)$ en $\mathbb{R}^3$?*

Solución Estos vectores son la imagen de la base canónica bajo la transformación lineal con matriz.

$$\begin{pmatrix} 1 & 2 & 0 \\ 1 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Por el teorema del cambio de variables, el volumen del conjunto imagen B es $|\det g|$ veces 1, el volumen de A (el cubo unidad en $\mathbb{R}^3$). Este determinante resulta ser 2, de modo que el volumen pedido es 2. ♦

9.3.3 Ejemplo *Evalúese* $\int_0^1 \int_0^1 (x^4 - y^4)\,dx\,dy$ *usando el cambio de variables* $u = x^2 - y^2$, $v = 2xy$.

Solución Se toma $B =]0, 1[\times]0, 1[$ y g es la transformación que aplica (u,v) en el (x, y) correspondiente, de modo que $g^{-1}(x, y) = (x^2 - y^2, 2xy)$. La región A correspondiente a B se muestra en la figura 9.3-2. Podemos verificar que $g : A \to B$ es inyectiva y que

$$Jg(u, v) = \frac{1}{Jg^{-1}(x, y)} = \frac{1}{4(x^2 + y^2)},$$

pues

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix} = 4(x^2 + y^2),$$

que no se anula. Así, por el teorema del cambio de variables, nuestra integral es

$$\int_A u \cdot (x^2 + y^2) \frac{du\,dv}{4(x^2 + y^2)} = \frac{1}{4} \int_A u\,du\,dv = \int_0^2 \int_{(v^2/4)-1}^{1-(v^2/4)} u\,du\,dv = 0.$$

(Estrictamente, se debería aplicar el teorema del cambio de variables primero sólo a la integral $\int_\varepsilon^1 \int_\varepsilon^1 (x^4 - y^4)\,dx\,dy$ y después hacer $\varepsilon \to 0$, para que $Jg(u, v)$ este acotado; un refinamiento muestra que esta hipótesis no es realmente necesaria.) La integral dada se puede calcular también directamente como una integral iterada:

$$\int_0^1 \int_0^1 (x^4 - y^4)\,dx\,dy = \int_0^1 \left(\frac{1}{5} - y^4\right) dy = 0. \quad ♦$$

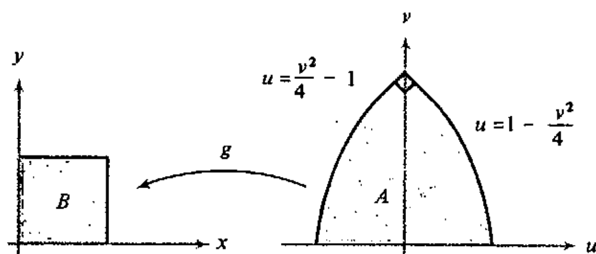


FIGURA 9.3-2 Un cambio de variables en $\mathbb{R}^2$

Ejercicios de §9.3

1. Muéstrese que el teorema del cambio de variables tiene como caso particular el teorema de una variable analizado en **§9.1.**

2. ¿Cuál es el volumen del paralelepípedo generado por (1, 1, 1, 1), (0, 1, 1, 0), (2, 0, 3, 0), (1, 1, 0, 1) en $\mathbb{R}^4$?

3. Transfórmese $\int_0^1 \int_y^1 (x^2 + y^2)\, dx\, dy$ usando el cambio de variables $x = u + v$, $y = u - v$.

4. Muéstrese que el volumen del paralelepípedo generado por los vectores $v_1, \ldots, v_n$ en $\mathbb{R}^n$ está dado por $|\det A_{ij}|^{1/2}$, donde $A_{ij} = \langle v_i, v_j \rangle$.

5. Muéstrese que el área de la elipse $x^2/a^2 + y^2/b^2 \leq 1$ es πab haciendo un cambio de variables y reduciendo el problema al de determinar el área de un círculo.

§9.4 Coordenadas polares

Una aplicación de la fórmula del cambio de variables es la evaluación de integrales por medio de las coordenadas polares. La función que pasa de coordenadas polares a coordenadas rectangulares es $g(r, \theta) = (r \cos \theta, r \operatorname{sen} \theta)$. El determinante jacobiano es

$$Jg(r,\theta) = \begin{vmatrix} \cos\theta & -r\operatorname{sen}\theta \\ \operatorname{sen}\theta & r\cos\theta \end{vmatrix} = r\cos^2\theta + r\operatorname{sen}^2\theta = r.$$

Si g está definida en el conjunto $\{(r, \theta) \mid r > 0, 0 < \theta < 2\pi\}$, entonces $Jg(r, \theta)$ nunca se anula y g es inyectiva en este conjunto. Dejaremos que el lector verifique que g es inyectiva. Aunque la imagen de esta función excluye el conjunto de puntos sobre el eje X con $x \geq 0$, éste es un conjunto de medida nula y, por lo tanto, no contribuye al valor de una integral (véase la figura 9.4-1).

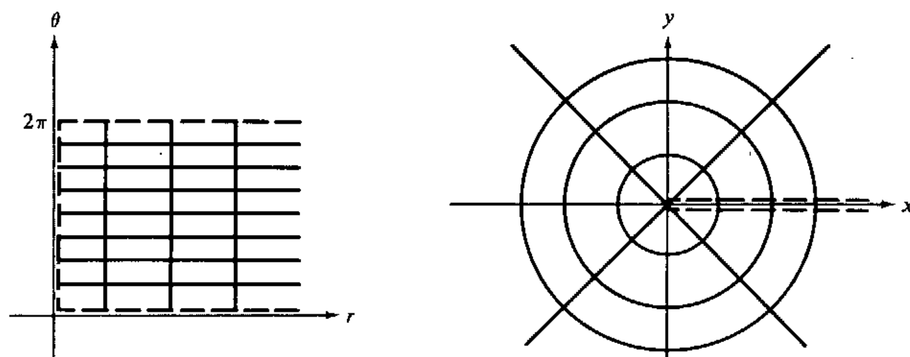


FIGURA 9.4-1 Coordenadas polares

9.4.1 Ejemplo *Determínese la masa de una placa delgada con la forma de un anillo, con radio interno* 1, *radio externo* 2 y *densidad de masa* $1/r^3$ *en los puntos que están a distancia r del centro. Véase la figura 9.4-2.*

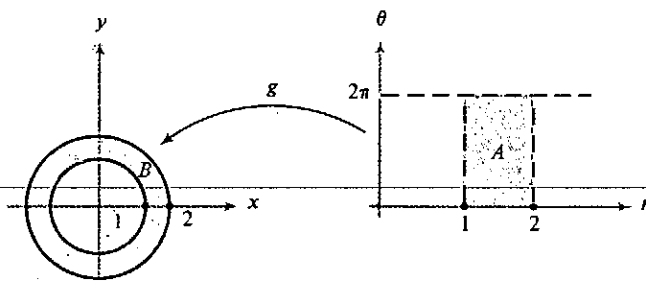


FIGURA 9.4-2 Descripción de un anillo en coordenadas polares

Solución Si B denota el anillo, entonces B es la imagen (salvo por los puntos en el eje X positivo) de $A = \{(r, \theta) \mid 0 < \theta < 2\pi,\ 1 < r < 2\}$ bajo la transformación de coordenadas polares. Entonces, la masa pedida está dada por

$$\int_B (x^2 + y^2)^{-3/2}\, dx\, dy = \int_A (1/r^3)|Jg|\, dr\, d\theta = \int_A 1/r^2\, dr\, d\theta$$
$$= \int_0^{2\pi}\int_1^2 1/r^2\, dr\, d\theta = \int_0^{2\pi}\left(-\frac{1}{2} + 1\right) d\theta = \pi. \quad \blacklozenge$$

Ejercicios de §9.4

Evalúense las integrales de los ejercicios 1 y 2 usando coordenadas polares.

1. $\int_D \exp(x^2 + y^2)\, dx\, dy$, donde $D = \{(x, y) \mid x^2 + y^2 < 1\}$.

2. $\int_D \ln(x^2 + y^2)\, dx\, dy$, donde $D = \{(x, y) \mid x \geq 0,\ y \geq 0,\ a^2 \leq x^2 + y^2 \leq b^2\}$ y donde a y b son constantes reales positivas.

3. Determínese el área de un círculo de radio r usando coordenadas polares.

§9.5 Coordenadas esféricas y cilíndricas

Las mismas técnicas que aplicamos a las coordenadas polares se pueden aplicar a las coordenadas esféricas. En este caso, sea $g(r, \varphi, \theta) = (r \operatorname{sen} \varphi \cos\theta,\ r \operatorname{sen}\theta \operatorname{sen}\varphi,\ r\cos\varphi)$ y considérese que g está definida en $\{(r, \varphi, \theta) \mid r > 0,\ 0 < \theta < 2\pi,\ 0 < \varphi < \pi\}$. Véase la figura 9.5-1. La imagen de este conjunto bajo g es todo $\mathbb{R}^3$, excepto la parte del plano xz tal que $x \geq 0$. Por el ejercicio 5 del capítulo 8, éste es un conjunto de medida nula, por lo que se puede despreciar en las integrales. El determinante jacobiano está dado por

$$Jg(r, \varphi, \theta) = \begin{vmatrix} \operatorname{sen}\varphi\cos\theta & r\cos\varphi\cos\theta & -r\operatorname{sen}\varphi\operatorname{sen}\theta \\ \operatorname{sen}\varphi\operatorname{sen}\theta & r\cos\varphi\operatorname{sen}\theta & r\operatorname{sen}\varphi\cos\theta \\ \cos\varphi & -r\operatorname{sen}\varphi & 0 \end{vmatrix} = r^2\operatorname{sen}\varphi.$$

Por lo tanto, $Jg(r, \varphi, \theta) \neq 0$ en la región dada; g también es inyectiva en esa región, por lo que la fórmula del cambio de variables implica

$$\int_{g(A)} f(x, y, z)\, dx\, dy\, dz = \int_A f(r\operatorname{sen}\varphi\cos\theta,\ r\operatorname{sen}\varphi\operatorname{sen}\theta,\ r\cos\varphi)\, r^2\operatorname{sen}\varphi\, dr\, d\varphi\, d\theta.$$

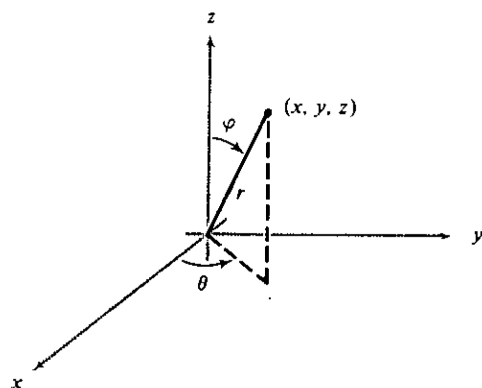


FIGURA 9.5-1 Coordenadas esféricas en $\mathbb{R}^3$

9.5.1 Ejemplo *Intégrese la función $f(x, y, z) = x^2 + y^2 + z^2$ sobre el conjunto $B = \{(x, y, z) \mid x^2 + y^2 + z^2 < 1\}$.*

Solución El conjunto B es la imagen de $A = (r, \varphi, \theta) \mid 0 < r < 1, 0 < \theta < 2\pi, 0 < \varphi < \pi\}$ bajo g (excepto los puntos de B en el plano xz, donde $x \geq 0$). En consecuencia,

$$\int_B (x^2 + y^2 + z^2)\, dx\, dy\, dz = \int_A r^2 |Jg|\, dr\, d\varphi\, d\theta = \int_A r^2 \cdot r^2 \operatorname{sen}\varphi\, dr\, d\varphi\, d\theta$$

(ya que sen $\varphi > 0$ en la región correspondiente). Así, nuestra integral es

$$\int_0^{2\pi} \int_0^{\pi} \int_0^1 r^4 \operatorname{sen}\varphi\, dr\, d\varphi\, d\theta = \int_0^{2\pi} \int_0^{\pi} \frac{\operatorname{sen}\varphi}{5}\, d\varphi\, d\theta = \frac{1}{5} \int_0^{2\pi} 2\, d\theta = \frac{4}{5}\pi. \quad \blacklozenge$$

Las coordenadas cilíndricas se tratan de forma análoga a las coordenadas polares y esféricas. La transformación adecuada es $g(r, \theta, z) = (r \cos\theta, r \operatorname{sen}\theta, z)$ sobre el conjunto $\{(r, \theta, z) \mid r > 0, 0 < \theta < 2\pi\}$. Véase la figura 9.5-2. El jacobiano es $Jg(r, \theta, z) = r$, de modo que el teorema del cambio de variables implica

$$\int_{g(A)} f(x, y, z)\, dx\, dy\, dz = \int_A f(r \cos\theta, r \operatorname{sen}\theta, z) r\, dr\, d\theta\, dz.$$

Esto es útil para integrales triples con "simetría cilíndrica", frente a los problemas con "simetría esférica", que se resuelven con coordenadas esféricas. ♦

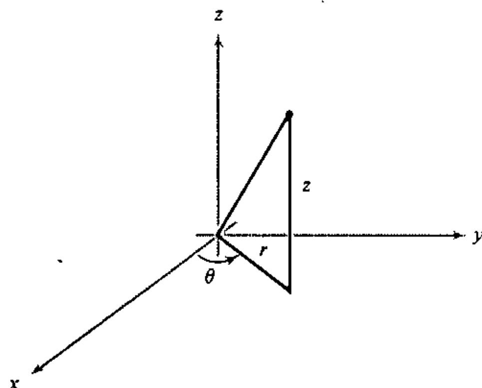


FIGURA 9.5-2 Coordenadas cilíndricas en $\mathbb{R}^3$

9.5.2 Ejemplo *Evalúese $\int_R ze^{-x^2-y^2}\,dx\,dy\,dz$ sobre la región $R = \{(x, y, z) \mid x^2 + y^2 \leq 1, 0 \leq z \leq 1$.*

Solución En este caso tenemos

$$\int_{z=0}^{1}\int_{\theta=0}^{2\pi}\int_{r=0}^{1} ze^{-r^2}\,r\,dr\,d\theta\,dz = \frac{\pi}{2}(1 - e^{-1}). \quad \blacklozenge$$

Ejercicios de §9.5

1. Evalúese $\int_D e^{(x^2+y^2+z^2)^{3/2}}$, donde D es la bola unidad en $\mathbb{R}^3$.

2. Sea D la bola unidad en $\mathbb{R}^3$. Evalúese $\int_D (dx\,dy\,dz)/\sqrt{1 + x^2 + y^2 + z^2}$.

3. Evalúese $\int_D z\sqrt{x^2 + y^2}\,dx\,dy\,dz$, donde $D = \{(x, y, z) \mid 1 \leq x^2 + y^2 \leq 2, 1 \leq z \leq 2\}$.

4. Hágase un cambio de variables para evaluar $\int_D \exp\,(x + y)/(x - y))\,dx\,dy$, donde $D = \{(x, y) \mid 0 \leq y \leq x, 0 \leq x \leq 1\}$.

5. Calcúlese

$$\int_0^1\int_{-1}^{1}\int_{-(1-y^2)^{1/2}}^{(1-y^2)^{1/2}} z(x^2 + y^2)\,dx\,dy\,dz$$

por medio de coordenadas cilíndricas.

§9.6 Un comentario sobre la integral de Lebesgue

En diversas partes de este y el anterior capítulo hemos dado indicios de la existencia de otra teoría de la integración. Tal teoría existe y se llama teoría de Lebesgue. Ahora analizaremos la necesidad de dicha teoría y sus diferencias básicas subyacentes con la teoría de Riemann.

La necesidad de la teoría de Lebesgue es, en gran medida, técnica: algunas funciones no son integrables Riemann, pero tal vez necesitemos integrarlas de algún modo. Por ejemplo, una función de este tipo puede aparecer como el límite de funciones integrables Riemann (un límite uniforme de funciones integrables Riemann es integrable Riemann, pero un límite puntual puede no serlo). Sin embargo, es deseable trabajar con espacios "completos", que contienen los límites de sucesiones de Cauchy, como, por ejemplo, $\mathbb{R}^n$ en el capítulo 2 y $C(A, \mathbb{R})$ en el capítulo 5. En el siguiente capítulo, que trata las series de Fourier, veremos un útil espacio de funciones, con la propiedad de que la convergencia en ese espacio puede ser puntual pero no necesariamente uniforme. La teoría de Riemann no basta para integrar tales funciones límite.

El problema de Henri Lebesgue era determinar una teoría de integración más general que la teoría de Riemann y que, además, fuera más útil desde el punto de vista técnico. En realidad, esto es una simplificación; la historia real es mucho más complicada. Por ejemplo, su trabajo recibió cierto reconocimiento por primera vez cuando usó sus ideas para dar una caracterización de las funciones integrables Riemann (véase el teorema **8.3.1**, capítulo 8). Para los detalles históricos, consúltese, por ejemplo, M. Kline,[1] *Mathematical Thought from Ancient to Modern Times* (Oxford University Press, Nueva York, 1972).

Daremos aquí un breve vistazo a estas ideas. Para desarrollarlas completamente se requiere todo un curso (véase, por ejemplo, H.L. Royden, *Real Analysis*, 3ª ed., MacMillan, Nueva York, 1988).

La base de la teoría de Lebesgue es la siguiente. En la teoría de Riemann, la integral se definía dividiendo el área por debajo de la gráfica de una función en rectángulos verticales. Pero, ¿por qué no dividirla también en rectángulos horizontales? Véase la figura 9.6-1.

Intuitivamente, este tipo de subdivisión en la que se toma una partición en el eje Y en vez de en el eje X tendría que dar la misma área. Pero existe, de hecho, una diferencia técnica, y la idea de Lebesgue (rectángulos horizontales) es precisamente la que queremos. Esto nos lleva también a unas matemáticas más complicadas; veamos por qué.

Sea $f : [a, b] \to \mathbb{R}$ y supóngase que f está acotada y que es no negativa. Sea $y_0 < y_1 < \cdots < y_n$ una partición de un intervalo que contiene a la imagen de f. Nuestro candidato como suma aproximante es (obsérvese la figura 9.6-1):

$$\sum (y_{i+1} - y_i) \cdot (\text{"longitud" del conjunto } \{x \mid f(x) \geq y_i\}).$$

[1] Traducción al castellano: M. Kline, *El pensamiento matemático de la antigüedad a nuestros días*, 3 vols., Alianza Editorial, 1992. (*N. del R.T.*)

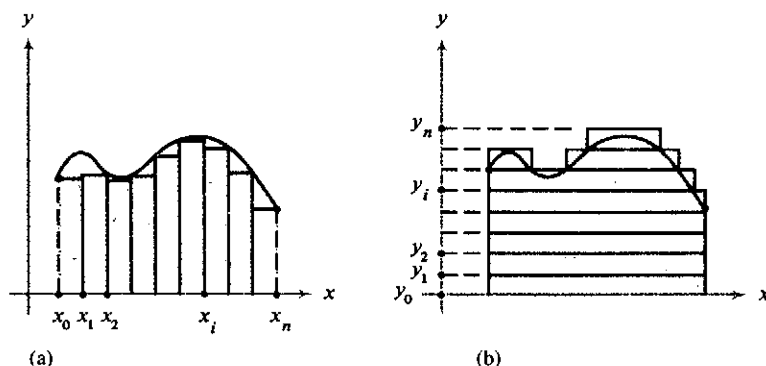


FIGURA 9.6-1 (a) Aproximaciones de Riemann; (b) aproximaciones de Lebesgue

El detalle técnico es que $\{x \mid f(x) \geq y_i\} = I_i$ podría ser un conjunto complicado; después de todo, la teoría de Riemann ya podía trabajar con funciones "decentes" y controlables, de modo que debemos estar preparados para trabajar con funciones un tanto complicadas. Así, nuestro problema es que si I_i es un conjunto complicado, ¿cómo calculamos su longitud? Aquí es donde entra una parte fundamental de la teoría de Lebesgue. Primero hay que desarrollar la idea de la medida (o longitud) de un conjunto. Esto se convierte en algo un tanto complicado, pero una vez desarrollada esta *teoría de la medida*, se puede proceder al estudio de las sumas aproximantes de Lebesgue en la forma que acabamos de describir. Obsérvese que ya hemos visto un aspecto de la "medida" en el capítulo 8, cuando estudiamos los conjuntos de medida nula. Este concepto está tomado directamente de la teoría de Lebesgue.

En consecuencia, la conclusión es que, con considerable más esfuerzo, es posible construir una teoría más amplia de la integración (que incluye la teoría de Riemann como caso particular). En áreas más avanzadas de las matemáticas, los logros técnicos valen más que el esfuerzo invertido.

§9.7 Operaciones de intercambio de límites

En el siguiente capítulo comenzaremos el estudio de un área muy vasta del análisis que ha sido la inspiración y una de las muchas aplicaciones de gran parte del análisis desarrollado en los dos últimos siglos, incluyendo una gran parte del contenido de este libro. Nos referimos al desarrollo de una función dada en términos de senos y cosenos, ya sea como una integral o como una suma, y se conoce como *análisis de Fourier*. Antes de esto, sería bueno hacer una comparación de varios de los resultados que hemos visto en el texto hasta ahora.

El análisis se ocupa principalmente de los procesos infinitos. Las cantidades analizadas son con frecuencia los valores de un límite de cierto tipo. Esto se aplica a las

derivadas, las integrales, las sumas de series infinitas, etcétera. Muchos de los teoremas que hemos encontrado pueden plantearse diciendo que, en ciertas condiciones, algunas operaciones de paso al límite se pueden intercambiar o realizar en un orden arbitrario.

9.7.1 Ejemplo *Un límite uniforme de funciones continuas es continuo.*

Uno de los primeros resultados del capítulo 5 era que si $f_1, f_2, f_3 \ldots$ son funciones continuas que convergen uniformemente a una función f, entonces f también es continua. La continuidad en un punto a expresa el hecho de que $\lim_{x\to a} f(x) = f(a)$, de modo que el resultado de este teorema se puede expresar como

$$\lim_{x\to a}\left(\lim_{n\to\infty}(f_n(x)\right) = \lim_{n\to\infty}\left(\lim_{x\to a}(f_n(x)\right)$$

La expresión de la izquierda es $\lim_{x\to a} f(x)$, mientras que la de la derecha es $\lim_{n\to\infty}(f_n(a)) = f(a)$.

9.7.2 Ejemplo *Bajo hipótesis adecuadas, la integral de un límite es el límite de las integrales.*

Este tipo de conclusión afirma también la validez del intercambio de dos operaciones de límite. En este caso estamos interesados en conclusiones de la forma

$$\int_a^b \left(\lim_{n\to\infty} f_n(x)\right) dx = \lim_{n\to\infty}\left(\int_a^b f_n(x)\,dx\right)$$

En el capítulo 5 vimos ejemplos que muestran que este intercambio no siempre se puede hacer. También vimos un caso bastante importante en el que sí se puede.

Si $f_1, f_2, f_3 \ldots$ es una sucesión de funciones integrables en un intervalo $[a, b]$ que convergen uniformemente en $[a, b]$ a f, entonces f es integrable y

$$\int_a^b f(x)\,dx = \lim_{n\to\infty}\left(\int_a^b f_n(x)\,dx\right).$$

9.7.3 Ejemplo *Operaciones término a término.*

Los teoremas que afirman la validez de la derivación o integración término a término de las series bajo diversas condiciones son también de este tipo. Sus conclusiones se pueden expresar como

$$\int_a^b \sum_{n=1}^{\infty} f_n(x)\,dx = \sum_{n=1}^{\infty} \int_a^b f_n(x)\,dx$$

o como

$$\frac{d}{dx}\left[\sum_{n=1}^{\infty} f_n(x)\right] = \sum_{n=1}^{\infty} \frac{d}{dx} f_n(x).$$

Cualquiera de estas conclusiones necesita ciertas condiciones para ser válidas. Encontramos algunos teoremas de este tipo en el capítulo 5, donde fueron particularmente eficaces en el caso de las series de potencias.

Derivación e integración

Existen varios resultados relativos a las integrales y las derivadas, algunos de los cuales ya hemos visto; recientemente, el teorema de Fubini, el cual afirma que las integrales iteradas se pueden evaluar en cualquier orden, con las condiciones adecuadas: se puede intercambiar un par de integraciones. Anteriormente, en el capítulo 6 hemos visto el resultado de que, con hipótesis bastante débiles, podemos concluir que las derivadas parciales cruzadas se pueden evaluar en cualquier orden: se puede intercambiar un par de operaciones de derivación. Es interesante ver la relación entre algunos de estos resultados, tal vez con hipótesis un poco más restrictivas. El propio teorema fundamental del cálculo afirma que se pueden intercambiar una operación de integración y una derivación.

Teorema *Si f y f' son continuas en $[a, b]$ y $a < x < b$, entonces*

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x) = f(a) + \int_a^x f'(t)\,dt$$

Este resultado es el núcleo de todo nuestro desarrollo posterior, pero comenzaremos con una versión un tanto distinta del intercambio de las operaciones de derivación e integración.

9.7.4 Ejemplo *Derivación de una integral que depende de un parámetro (derivación bajo el signo de integral). En condiciones razonables,*

$$\frac{d}{dy}\left(\int_a^b f(x,y)\,dx\right) = \int_a^b \frac{\partial f}{\partial y}(x,y)\,dx$$

He aquí una versión precisa:

Proposición *Supóngase que $f : [a, b] \times [c, d] \to \mathbb{R}$ es continua y que $\partial f/\partial y$ existe para $c < y < d$ y se extiende de forma continua a $[a, b] \times [c, d]$. Sea*

$$F(y) = \int_a^b f(x, y)\, dx.$$

Entonces F es derivable y

$$F'(y) = \int_a^b \frac{\partial f}{\partial y}(x, y)\, dx.$$

Demostración Necesitamos mostrar que el límite del cociente de diferencias existe y que es igual a la integral deseada. Así, considérese

$$\left| \frac{F(y+h) - F(y)}{h} - \int_a^b \frac{\partial f}{\partial y}(x, y)\, dx \right| = \left| \int_a^b \left(\frac{f(x, y+h) - f(x, y)}{h} - \frac{\partial f}{\partial y}(x, y) \right) dx \right|.$$

Por el teorema del valor medio,

$$\frac{f(x, y+h) - f(x, y)}{h} = \frac{\partial f}{\partial y}(x, c_{x,h})$$

para algún $c_{x,h}$ entre y e $y + h$, de modo que

$$\left| \frac{F(y+h) - F(y)}{h} - \int_a^b \frac{\partial f}{\partial y}(x, y)\, dx \right| = \left| \int_a^b \left(\frac{\partial f}{\partial y}(x, c_{x,h}) - \frac{\partial f}{\partial y}(x, y) \right) dx \right|.$$

Como $\partial f/\partial y$ se extiende de forma continua en el conjunto compacto $[a, b] \times [c, d]$, es uniformemente continua. Dado $\varepsilon > 0$, podemos elegir $\delta > 0$ tal que

$$\left| \frac{\partial f}{\partial y}(x_0, y_0) - \frac{\partial f}{\partial y}(x, y) \right| < \frac{\varepsilon}{b - a}$$

siempre que $|x - x_0| < \delta$ y $|y - y_0| < \delta$. Si $|h| < \delta$, entonces

$$\left|\frac{F(y+h) - F(y)}{h} - \int_a^b \frac{\partial f}{\partial y}(x,y)\,dx\right| \le \int_a^b \left|\frac{\partial f}{\partial y}(x, c_{x,h}) - \frac{\partial f}{\partial y}(x,y)\right| dx$$

$$< \frac{\varepsilon}{b-a}(b-a) = \varepsilon.$$

Esto establece el límite deseado. ■

Existen resultados similares para las integrales impropias, pero requieren algún tipo de uniformidad. Esta situación es análoga a la de las sumas, aunque esto no es sorprendente si consideramos la integral como un límite de sumas. La derivada de la suma de un número finito de términos es igual a la suma de las derivadas. La misma conclusión para series infinitas requiere de ciertas hipótesis. El teorema que obtuvimos para series utilizaba la convergencia uniforme de la serie de derivadas. He aquí una hipótesis similar acerca de la integral de la derivada parcial.

9.7.5 Proposición *Supóngase que $f: [a, \infty) \times [c, d] \to \mathbb{R}$ es continua y, como en la proposición anterior, $\partial f/\partial y$ existe y es continua en $[a, \infty] \times [c, d]$. Supóngase además que las integrales*

y $$\int_a^\infty f(x,y)\,dx \quad \text{y} \quad \int_a^\infty \frac{\partial f}{\partial y}(x,y)\,dx$$

convergen, la segunda uniformemente con respecto a $y \in [c, d]$. Sea

$$F(y) = \int_a^\infty f(x,y)\,dx.$$

Entonces F es derivable y

$$F'(y) = \int_a^\infty \frac{\partial f}{\partial y}(x,y)\,dx.$$

La hipótesis de la convergencia uniforme de la segunda integral a algún $G(y)$ significa que para cada $\varepsilon > 0$ existe B tal que

$$\left|G(y) - \int_a^b \frac{\partial f}{\partial y}(x,y)\,dx\right| < \varepsilon \text{ para todo } y \in [c,d] \text{ siempre que } b > B.$$

La demostración es análoga a la del resultado correspondiente acerca de la derivación término a término de las series.

Estos dos resultados justifican la *derivación bajo el signo de integral* en muchos casos. Esta idea tiene un valor práctico importante en los cálculos; en los ejercicios aparecen algunos ejemplos. En el ejemplo resuelto **9.2** al final del capítulo, incluimos este resultado en una fórmula general para derivadas de integrales. Aquí estamos interesados en la forma en que esto se relaciona con otros teoremas de intercambio, como el de Fubini.

9.7.6 Ejemplo *Teorema de Fubini: Con las condiciones adecuadas, las integrales iteradas se pueden evaluar en cualquier orden:*

$$\int_a^b \left[\int_c^d f(x,y)\,dy\right] dx = \int_c^d \left[\int_a^b f(x,y)\,dx\right] dy.$$

Es interesante ver cómo una demostración para f continua se puede basar en el proceso de derivación bajo el signo de la integral que acabamos de obtener.

La función $h(t, y) = \int_a^t f(x, y)\,dx$ es continua en $[a, b] \times [c, d]$, al igual que su derivada parcial $\partial h/\partial t = f(t, y)$. La integral

$$g(x) = \int_c^d f(x,y)\,dy$$

es una función continua de x. Ahora sólo necesitamos considerar la función

$$H(t) = \int_a^t \left[\int_c^d f(x,y)\,dy\right] dx - \int_c^d \left[\int_a^t f(x,y)\,dx\right] dy.$$

Está claro que $H(a) = 0$ y queremos mostrar que $H(b) = 0$. Pero

$$H(t) = \int_a^t g(x)\,dx - \int_c^d h(t,y)\,dy$$

de modo que

$$H'(t) = g(t) - \int_c^d \frac{\partial h}{\partial t}\,dy = \int_c^d f(t,y)\,dy - \int_c^d f(t,y)\,dy = 0$$

para todo t. Así, H debe ser constante en $[a, b]$, de modo que $H(b) = 0$, como se deseaba.

Por último, veremos la forma de obtener la igualdad de las derivadas parciales cruzadas a partir del teorema de Fubini.

9.7.7 Ejemplo *Igualdad de las derivadas parciales cruzadas.*

Si $\dfrac{\partial^2 g}{\partial x \partial y}$ *y* $\dfrac{\partial^2 g}{\partial y \partial x}$ *existen y son continuas, entonces son iguales.*

Queremos mostrar que

$$\frac{\partial^2 g}{\partial x \partial y} - \frac{\partial^2 g}{\partial y \partial x} = 0.$$

Como es una función continua, basta mostrar que la integral sobre cualquier rectángulo pequeño $[a, b] \times [c, d]$ es 0. Si R es uno de tales rectángulos, entonces

$$\begin{aligned}
&\iint_R \left[\frac{\partial^2 g}{\partial x \partial y} - \frac{\partial^2 g}{\partial y \partial x}\right] dx\,dy \\
&= \int_c^d \left[\int_a^b \frac{\partial}{\partial x}\left(\frac{\partial g}{\partial y}\right) dx\right] dy - \int_a^b \left[\int_c^d \frac{\partial}{\partial y}\left(\frac{\partial g}{\partial x}\right) dy\right] dx \\
&= \int_c^d \left[\frac{\partial g}{\partial y}(b, y) - \frac{\partial g}{\partial y}(a, y)\right] dy - \int_a^b \left[\frac{\partial g}{\partial x}(x, d) - \frac{\partial g}{\partial x}(x, c)\right] dx \\
&= (g((b, d) - g(b, c) - g(a, d) + g(a, c)) - (g(b, d) - g(a, d) - g(b, c) + g(a, c)) \\
&= 0
\end{aligned}$$

como se deseaba. ■

Ejercicios de §9.7

1. **a.** Sea $t > -1$. Muéstrese que $\int_0^1 x^t \log(x)\, dx$ existe.

b. Evalúese la integral en **a** considerando

$$\frac{d}{dt}\left(\int_0^1 x^t\, dx\right)$$

c. Sea n un entero positivo y $t > -1$. Evalúese $\int_0^1 x^t\, (\log(x))^n\, dx$.

2. Muéstrese que

$$\int_0^{2\pi} x \operatorname{sen}(tx)\, dx = \frac{\operatorname{sen}(2\pi t) - 2\pi t \cos(2\pi t)}{t^2}$$

considerando la derivada de $\int_0^{2\pi} \cos(tx)\, dx$ con respecto a t.

3. Evalúese $\dfrac{d}{dt}\displaystyle\int_0^{2\pi} x \cos(tx)\, dx$.

4. Demuéstrese la proposición **9.7.5.**

5. Evalúese $\int_0^\infty x^n\, e^{-ax}\, dx$ derivando varias veces $\int_0^\infty e^{-ax}\, dx$ (aquí $a > 0$). Justifíquese la aplicación de la proposición **9.7.5.**

Demostraciones de los teoremas del capítulo 9

9.2.1 Teorema

i. *Sea A el rectángulo descrito por $a \le x \le b$, $c \le y \le d$ y sea $f: A \to \mathbb{R}$ continua. Entonces*

$$\int_A f = \int_a^b \left(\int_c^d f(x,y)\,dy \right) dx = \int_c^d \left(\int_a^b f(x,y)\,dx \right) dy.$$

La expresión

$$\int_a^b \left(\int_c^d f(x,y)\,dy \right) dx$$

significa que la función

$$g(x) = \int_c^d f(x,y)\,dy$$

se integra de a a b.

ii. *En* **i**, *supóngase que f es integrable y que la función $f_x: [c, d] \to \mathbb{R}$ definida por $f_x(y) = f(x, y)$ es integrable para cada $x \in [a, b]$ fijo. Entonces*

$$\int_A f = \int_a^b \left(\int_c^d f(x,y)\,dy \right) dx.$$

De manera análoga, podemos suponer que

$$\int_a^b f(x,y)\,dx$$

existe para cada y, y obtener

$$\int_A f = \int_c^d \left(\int_a^b f(x,y)\,dx \right) dy.$$

Demostración Como **i** es un caso particular de **ii**, sólo necesitamos demostrar **ii.** Sea $g: [a, b] \subset \mathbb{R} \to \mathbb{R}$ la función dada por

$$g(x) = \int_c^d f(x,y)\,dy.$$

Debemos mostrar que g es integrable en $[a, b]$ y que $\int_A f = \int_a^b g(x)\, dx$.

Supóngase que $a = x_0 < x_1 < \cdots < x_n = b$ y $c = y_0 < \cdots < y_n = d$ son particiones de los intervalos $[a, b]$ y $[c, d]$. Sea $P_{[a,b]}$ la partición de $[a, b]$ dada por los conjuntos $V_i = [x_{i-1}, x_i]$, $P_{[c,d]}$ la partición de $[c, d]$ dada por los conjuntos $W_j = [y_{j-1}, y_j]$ y P_A la partición de A dada por los conjuntos

$$S_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j].$$

Entonces

$$L(f, P_A) = \sum_{i,j} m_{S_{ij}}(f)v(S_{ij}) = \sum_i \sum_j m_{S_{ij}}(f)v(W_j)v(V_i),$$

donde $m_S(f)$ es el ínfimo mínimo, por la continuidad de f sobre el conjunto S. Para $x \in V_i$, tenemos $m_{s_{ij}}(f) \leq m_{w_j}(f_x)$, donde f_x está definida como $f_x(y) = f(x, y)$. Por lo tanto,

$$\sum_j m_{S_{ij}}(f)v(W_j) \leq \sum_j m_{W_j}(f_x)v(W_j) \leq \int_c^d f_x(y)\, dy = g(x).$$

Como esta desigualdad es válida para cada $x \in V_i$, obtenemos $\sum_j m_{s_{ij}}(f)v(W_j) \leq m_{v_i}(g)$. En consecuencia,

$$L(f, P_A) \leq \sum_i m_{V_i}(g)v(V_i) \leq L(g, P_{[a,b]}).$$

A partir de esto y de un argumento similar para las sumas superiores, obtenemos las desigualdades

$$L(f, P_A) \leq L(g, P_{[a,b]}) \leq U(g, P_{[a,b]}) \leq U(f, P_A).$$

Como f es integrable en A, estas desigualdades muestran que g es integrable y

$$\int_a^b \left(\int_c^d f(x, y)\, dy \right) dx = \int_a^b g(x)\, dx = \int_A f.$$

Con el mismo argumento, si suponemos que $\int_a^b f(x, y)\, dx$ existe para cada y, obtenemos

$$\int_A f = \int_c^d \left(\int_a^b f(x, y)\, dx \right) dy. \quad \blacksquare$$

9.2.2 Corolario *Sean $\varphi, \psi : [a, b] \to \mathbb{R}$ transformaciones continuas tales que $\varphi(x) \leq \psi(x)$ para todo $x \in [a, b]$; sea $A = \{(x, y) \mid a \leq x \leq b,\ \varphi(x) \leq y \leq \psi(x)\}$, y sea f: $A \to \mathbb{R}$ continua. Entonces (véase la figura 9.2-1)*

$$\int_A f = \int_a^b \left(\int_{\varphi(x)}^{\psi(x)} f(x, y)\, dy \right) dx.$$

Demostración Sea $S = [a, b] \times [c, d]$ un rectángulo cerrado que contenga a A y extendamos f a S haciéndola igual a cero en $S \backslash A$. Los conjuntos gráfica(φ) $= \{x, \varphi(x)) \mid x \in [a, b]\}$ y gráfica(ψ) $= \{(x, \psi(x)) \mid x \in [a, b]\}$ tienen medida nula (por el ejercicio 23 del capítulo 8). Así, el conjunto de discontinuidades de f definida en S tiene medida nula, por lo que f es integrable en S. Además, para cualquier x, f_x es continua en $[c, d]$, excepto tal vez en $\varphi(x)$ y $\psi(x)$, de modo que f_x es integrable para cualquier $x \in [a, b]$. Aplicando el teorema **9.2.1** obtenemos

$$\int_A f = \int_S f = \int_a^b \int_c^d f_x(y)\, dy\, dx = \int_a^b \left(\int_{\varphi(x)}^{\psi(x)} f_x(y)\, dy \right) dx. \quad \blacksquare$$

La demostración del teorema **9.2.3** es análoga a la demostración de **9.2.2**, de modo que la dejamos como ejercicio.

Pasemos ahora al teorema **9.3.1**. Su demostración puede ser muy laboriosa si no se lleva a cabo de una manera eficiente. En particular, la idea dada en el texto es difícil de precisar si la interpretamos de un modo demasiado literal. La demostración que vamos a dar es de J. Schwartz, *American Mathematical Monthly* **61** (1954), págs. 81–85.

9.3.1 Teorema del cambio de variables *Sean $A \subset \mathbb{R}^n$ un conjunto abierto acotado con volumen y sea $g : A \to \mathbb{R}^n$ una transformación C^1 inyectiva. Supóngase que $Jg(x) \neq 0$ para todo $x \in A$ y que $|Jg(x)|$ y $1/|Jg(x)|$ están acotados en A. Sea $B = g(A)$ y supóngase que B tiene volumen (véase* §**8.2**). *Si $f : B \to \mathbb{R}$ está acotada y es integrable, entonces $(f \circ g) \mid J(g)$ es integrable en A y*

$$\int_B f = \int_A (f \circ g)|Jg|,$$

es decir,

$$\int_B f(y_1, \ldots, y_n) dy_1 \cdots dy_n = \int_A f(g(x_1, \ldots, x_n)) \frac{\partial(g_1, \ldots, g_n)}{\partial(x_1, \ldots, x_n)}\, dx_1 \cdots dx_n.$$

Nota. Un análisis de la demostración muestra que f, $|Jg|$ y $|Jg|^{-1}$ no tienen que ser acotados (véase **§8.7**, para el caso de integrales impropias, así como la observación después del ejemplo resuelto **9.4**).

El primer paso de la demostración es el de establecer nuestra fórmula cuando $g = L$ es una transformación lineal, en cuyo caso $JL = \det L$. Esto proporciona la interpretación geométrica de $\det L$: *es el factor de cambio de los volúmenes bajo la transformación L.*

Como no queremos suponer esto del álgebra lineal, desarrollaremos la demostración con algo de detalle. Sin embargo, sí necesitamos recordar dos puntos del álgebra lineal: **(i)** $\det TS = \det T \cdot \det S$ y **(ii)** cualquier matriz es producto de matrices elementales.

Lema 1 *Si $L: \mathbb{R}^n \to \mathbb{R}^n$ es una transformación lineal y $A \subset \mathbb{R}^n$ es un conjunto que tiene volumen (es decir, $\int_A 1_A$ existe), entonces $L(A)$ tiene volumen y el volumen es $|\det L| \cdot v(A)$, esto es, $\int_{L(A)} 1_A = \int_A |\det L|$ (véase la figura 9.P-1).*

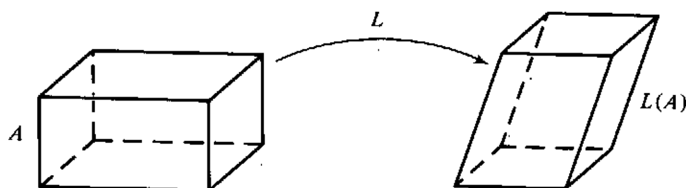


FIGURA 9.P-1 Imagen de un rectángulo bajo una transformación lineal

Demostración Primero consideramos el caso particular en que A es un rectángulo y L es una transformación lineal cuya matriz expresada en la base canónica es de uno de los dos tipos siguientes:

$$L_1 = \begin{pmatrix} 1 & 0 & & \cdots & & & 0 \\ 0 & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ \vdots & & & & c & & \vdots \\ & & & & & \ddots & \\ & & & & & 1 & 0 \\ 0 & & & \cdots & & 0 & 1 \end{pmatrix}$$

o

$$L_2 = \begin{pmatrix} 1 & 0 & & \cdots & & & 0 & \\ 0 & 1 & & & & & & \\ & & \ddots & & & & 1 & \\ & & & 1 & & & & \\ \vdots & & & & 1 & & & \ddots \\ & & & & & \ddots & & \\ & & & & & & 1 & 0 \\ 0 & & & \cdots & & & 0 & 1 \end{pmatrix}$$

La primera matriz se obtiene de la matriz identidad reemplazando un 1 de la diagonal por una constante c. La segunda matriz se obtiene de la matriz identidad escribiendo un 1 en cualquier parte fuera de la diagonal. Éstas son las ***matrices elementales***.

Si $A = [a_1, b_1] \times \cdots \times [a_n, b_n]$ y c está en la i-ésima fila, entonces

$$L_1(A) = [a_1, b_1] \times \cdots \times [ca_i, cb_i] \times \cdots \times [a_n, b_n],$$

de modo que el volumen de $L_1(A)$ es

$$v(L_1(A)) = |c|v(A) = |\det\ L_1|v(A).$$

Si el 1 fuera de la diagonal está en la posición (i, j), entonces

$$L_2(A) = \{(x_1, \ldots, x_{i-1}, x_i + x_j, x_{i+1}, \ldots, x_n) \mid x_k \in [a_k, b_k]\}$$

para $1 \leq k \leq n$. Esta imagen es un paralelepípedo en el que las secciones paralelas al plano $x_i\, x_j$ toman la forma de paralelogramos. Las secciones en todos los planos perpendiculares siguen siendo rectángulos. Las posibilidades aparecen en la figura 9.P-2, donde se muestran las secciones paralelas al plano $x_i\, x_j$.

Para nuestra conveniencia, renumeramos las coordenadas de modo que el "1" aparezca en la segunda columna de la fila superior, para que las variables afectadas sean x_1 y x_2. Geométricamente, vemos que el volumen no se ha alterado, pues estas secciones transversales son paralelogramos con la misma base y altura que el rectángulo original de dos dimensiones. Esto se puede confirmar analíticamente, examinando la integral que da el volumen de la imagen y usando el teorema de Fubini para que la integral con respecto a x_1 sea la penúltima y la integral con respecto a x_2 la última (continuamos con la numeración en la cual las variables afectadas por el corte son x_1 y x_2).

Así, el volumen de $L_2(A)$ es

$$\begin{aligned} v(L_2(A)) &= \int\int \cdots \int\int 1\, dx_2\, dx_1\, dx_3 \cdots dx_{n-1}\, dx_n \\ &= (b_n - a_n)(b_n - a_n) \times \cdots \times (b_3 - a_3) \int\int 1\, dx_2\, dx_1. \end{aligned}$$

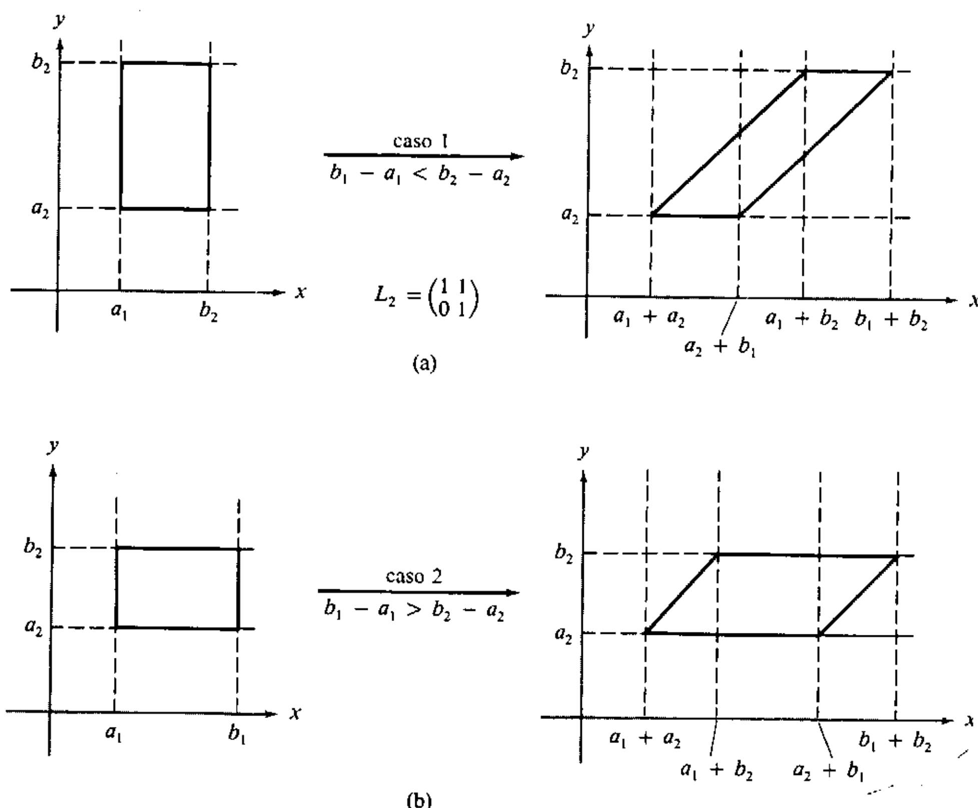


FIGURA 9.P-2 Imagen de un rectángulo bajo una matriz elemental L_2

Para cada x_2 entre a_2 y b_2, x_1 está entre las rectas de 45° $x = y + a_1$ y $x = y + b_1$. En consecuencia,

$$\int\int 1\, dx_1\, dx_2 = \int_{a_2}^{b_2} \left(\int_{x_2+a_1}^{x_2+b_1} 1\, dx_1 \right) dx_2 = \int_{a_2}^{b_2} (b_1 - a_1) dx_2$$
$$= (b_2 - a_2)(b_1 - a_1),$$

lo que lleva a que

$$v(L_2(A)) = (b_n - a_n)(b_n - a_n) \times \cdots \times (b_3 - a_3)(b_2 - a_2)(b_1 - a_1) = v(A).$$

Como $\det(L_2) = 1$, tenemos $v(L_2(A)) = |\det L_2|\, v(A)$.

Sea A un conjunto arbitrario con volumen y L_i una de las matrices elementales. Supóngase que $\det L_i \neq 0$ (es decir, $c \neq 0$ si L_i es una matriz elemental del primer tipo).

Sea S un rectángulo que contiene a A y sea P una partición de A en subrectángulos $S_1, \ldots,$ S_N de modo que

$$U(1_A, P) - v(A) < \varepsilon(2|\det\ L_i|) \quad \text{y} \quad v(A) - L_i(1_A, P) < \varepsilon(2|\det\ L|).$$

Si $V_\varepsilon = \bigcup\{S_i \mid S_i \subset A\}$ y $W_\varepsilon = \bigcup\{S_i \mid S_i \cap A \neq \varnothing\}$, obtenemos $v(L_i(V_\varepsilon)) = |\det L_i|\ (1_A, P)$ y $v(L_i(W_\varepsilon)) = |\det L_i|\ U(1_A, P)$, de modo que $v(L_i(W_\varepsilon)) - v(L_i(V_\varepsilon)) < \varepsilon$, y así $L_i(A)$ tiene volumen y $v(L_i(A)) = |\det L_i|\ v(A)$. Si $\det L_i = 0$, es decir, L_i es una matriz del primer tipo con $c = 0$, entonces $v(L_i(S)) = 0$ para cualquier rectángulo S, y lo mismo $v(L_i(A)) = 0$ para cualquier conjunto A con volumen.

Si L es una transformación lineal y A es un conjunto con volumen, entonces, del hecho de que cualquier matriz es el producto de matrices elementales concluimos que $L = L_1\ L_2 \cdots L_k$, donde las L_i son matrices elementales. Aplicando repetidamente lo que acabamos de demostrar se sigue que $L(A)$ tiene volumen y que

$$v(L(A)) = |\det\ L_1||\det\ L_2| \cdots |\det\ L_k| v(A) = |\det\ L| v(A). \quad \blacksquare$$

Lema 2 *Si el teorema es cierto para la función $f = 1$, entonces también es cierto para cualquier función integrable f.*

Demostración Si el teorema es cierto para $f = 1$, entonces es cierto para cualquier función constante (¿por qué?). Si f es cualquier función integrable en $g(A)$, sea S un rectángulo que contiene a $g(A)$ y sea P una partición de $g(A)$ en rectángulos $S_1, \ldots, S_N$. Sea $m_{S_i}(f) = \inf\{f(x) \mid x \in S_i\}$, lo que denotará también la función constante en S_i con valor constante $m_{S_i}(f)$. Entonces

$$\begin{aligned} L(f, P) &= \sum_{i=1}^{N} m_{S_i}(f) v(S_i) = \sum_{i=1}^{N} \int_{S_i} m_{S_i}(f) \\ &= \sum_{i=1}^{N} \int_{g^{-1}(S_i)} (m_{S_i}(f) \circ g)|Jg| \\ &\leq \sum_{i=1}^{N} \int_{g^{-1}(S_i)} (f \circ g)|Jg| \\ &= \int_{g^{-1}(S_i)} (f \circ g)|Jg| = \int_A (f \circ g)|Jg|. \end{aligned}$$

Por lo tanto, $\int_{g(A)} f \leq \int_A (f \circ g)\ |Jg|$. Un resultado análogo con $M_{S_i}(f) = \sup\{f(x) \mid x \in S_i\}$ muestra que $\int_{g(A)} f \geq \int_A (f \circ g)\ |Jg|$, de modo que $\int_{g(A)} f = \int_A (f \circ g)\ |Jg|$. $\blacksquare$

Lema 3 *El teorema es cierto si g es una transformación lineal.*

Demostración Por el lema **1**,

$$\int_{g(A)} 1 = \int_A |\det\ g| = \int_A |Jg|,$$

pues $g = \mathbf{D}g$. Por el lema **2**,

$$\int_{g(A)} f = \int_A (f \circ g)|Jg|. \quad \blacksquare$$

Necesitamos otra observación fundamental:

Lema 4 *Si el teorema es cierto para $g : A \to \mathbb{R}^n$ y para $h : B \to \mathbb{R}^n$, donde $g(A) \subset B$, entonces el teorema es válido para $h \circ g : A \to \mathbb{R}^n$.*

Demostración Realizamos el siguiente cálculo:

$$\begin{aligned}\int_{h\circ g(A)} f &= \int_{h(g(A))} f = \int_{g(A)} (f \circ h)|Jh| \\ &= \int_A (f \circ h \circ g)(|Jh| \circ g)|Jg| = \int_A (f \circ (h \circ g))|J(h \circ g)|. \quad \blacksquare\end{aligned}$$

Un poco más de notación nos permitirá desarrollar nuestra demostración de una forma sencilla. Si $x \in \mathbb{R}^n$, $x = (x_1, \ldots, x_n)$, sea $|x| = \text{máx}_{1\le i\le n} |x_i|$. Esta "norma" tiene la propiedad de que, en términos de ella, un cubo con centro en p y lado de longitud $2s$ se puede caracterizar como $|x-p| \le s$. Si $A : \mathbb{R}^n \to \mathbb{R}^n$ es la transformación lineal con matriz a_{ij}, de modo que

$$A(x) = A(x_1, \ldots, x_n) = \left(\sum_{j=1}^{n} a_{1j}x_j, \ldots, \sum_{j=1}^{n} a_{nj}x_j\right),$$

definimos

$$|A| = \underset{1\le i\le n}{\text{máx}} \sum_{j=1}^{n} |a_{ij}|.$$

Así, $|A(x)| \le |A||x|$. También utilizamos la matriz jacobiana $j(x) = (j_{ik}(x))$ de la transformación $g(x) = (g_1(x) \cdots g_n(x))$, donde, como es habitual, $j_{ik}(x) = \partial g_i(x)/\partial x_k$.

Si C es un cubo en el conjunto abierto A, tal que C es el conjunto de los x que satisfacen $|x-p| \leq s$, entonces $v(C) = (2s)^2$. Por el teorema del valor medio,

$$g_i(x) - g_i(p) = \sum_{k=1}^{n} j_{ik}[p + \theta_i(x)(x-p)](x_k - p_k),$$

donde $0 \leq \theta_i(x) \leq 1$. Así, $|g(x) - g(p)| \leq s \text{ máx}_{y\in C} |j(y)|$; es decir, $g(C)$ está totalmente contenido en el cubo definido por $|z - g(p)| \leq s \text{ máx}_{y\in C} |j(y)|$, de modo que si $g(C)$ tiene volumen, entonces

$$v(g(C)) \leq \left\{ \max_{y\in C} |j(y)|^n v(C) \right\}. \qquad (1)$$

Para garantizar que $g(C)$ tiene volumen, demostramos otro lema.

Lema 5 *Si $h : U \subset \mathbb{R}^n \to \mathbb{R}^n$ es una transformación C^1 inyectiva, $Jh(x) \neq 0$, U es un conjunto abierto acotado y C es un conjunto con volumen tal que* $\text{cl}(C) \subset U$, *entonces $h(C)$ tiene volumen.*

Demostración Basta mostrar que la frontera $\partial(h(C))$ tiene contenido nulo. Para ello, primero mostramos que $\partial(h(C)) \subset h(\partial(C))$. Sea $x \in \partial(h(C))$; para mostrar que $x \in h(\partial(C))$, sea $y = h^{-1}(x)$. Entonces debemos mostrar que $y \in \partial(C)$. Sea V una vecindad de y, y supóngase que $V \subset U$. Entonces $h(V)$ es una vecindad abierta de x, ya que h^{-1} es continua. Así, $h(V)$ contiene puntos de $h(C)$ y de $\mathbb{R}^n \backslash h(C)$, pues $x \in \partial(C)$. Como h es inyectiva, V contiene puntos de A y de $\mathbb{R}^n \backslash C$, de modo que $y \in \partial(C)$. Aplicando este argumento a h^{-1}, vemos que, de hecho, $\partial(h(C)) = h(\partial(C))$.

Para mostrar que $h(\partial(C))$ tiene volumen nulo, dado $\varepsilon > 0$, recubrimos $\partial(C)$ con rectángulos $B_1, \ldots, B_N$ de volumen total $\leq \varepsilon$. La ecuación (1) mostraba que $h(\partial(C))$ está contenido en un recubrimiento mediante rectángulos con volumen total $(\text{máx } |Jh(x)|)\varepsilon$, donde el máximo se evalúa sobre $B_1 \cup \cdots \cup B_N$. Así, $h(\partial(C))$ tiene volumen nulo. ■

Si A es una transformación lineal y S tiene volumen, entonces

$$v(A^{-1}(S)) = \det(A^{-1})v(S)$$

(tómese $f = 1$ en $A^{-1}(S)$, $f = 0$ en el complementario y aplíquese el lema 3). En (1), sea $S = g(C)$, que ahora sabemos que tiene volumen; como

$$|\det(A^{-1})| v(g(C)) \leq \left\{ \max_{y\in C} |A^{-1}j(y)| \right\}^n v(C),$$

obtenemos

$$g(C) \leq |\det(A)| \left\{ \max_{y \in C} |A^{-1}j(y)| \right\}^n v(C). \tag{2}$$

Subdividamos el cubo C en un conjunto finito $C_1, \ldots, C_M$ de cubos que no se solapen, centrados en $x_1, \ldots, x_M$ y supóngase que δ es mayor que la longitud de un lado de cualquiera de ellos. Aplicamos (2) a cada $C_1, \ldots, C_M$, tomando $A = j(x_i)$ en (2), y después sumamos, para obtener

$$v(g(C)) \leq \sum_{i=1}^{M} |\det(j(x_i))| \left\{ \max_{y \in C_i} |j^{-1}(x_i)j(y)| \right\}^n v(C_i).$$

Como $j(x)$ es una función continua (cuya imagen son matrices), $j^{-1}(z)j(y)$ tiende a la matriz identidad cuando $z \to y$, y por lo tanto $\{\max_{y \in C_i} |j^{-1}(x_i)j(y)|\}^n \leq 1 + \eta(\delta)$, donde $\eta(\delta)$ tiende a cero con δ. Así,

$$v(g(C)) \leq [1 + \eta(\delta)] \sum_{i=1}^{M} |\det(j(x_i))| v(C_i);$$

cuando δ tiende a cero, la suma en el miembro derecho tiende a $\int_C |Jg(x)|\, dx$, y la desigualdad queda

$$v(g(C)) \leq \int_C |Jg(x)|\, dx. \tag{3}$$

La demostración del lema **2** con $f = 0$ fuera de B, junto con (3), produce

$$\int_{g(A)} f \leq \int_A (f \circ g)|Jg|. \tag{4}$$

La ecuación (4) se puede aplicar también a g^{-1} para dar

$$\int_A (f \circ g)|Jg| \leq \int_{g(A)} (f \circ g \circ g^{-1})|Jg \circ g^{-1}| \cdot |Jg^{-1}|,$$

es decir,

$$\int_A (f \circ g)|Jg| \leq \int_{g(A)} f. \tag{5}$$

Combinando (4) y (5) se obtiene el teorema. ■

Ejemplos resueltos del capítulo 9

Ejemplo 9.1 *Úsese la fórmula del cambio de variables y las coordenadas polares para mostrar que*

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$$

(*la función* e^{-x^2} *se denomina* ***función gaussiana****; véase la figura 9.ER-1*).

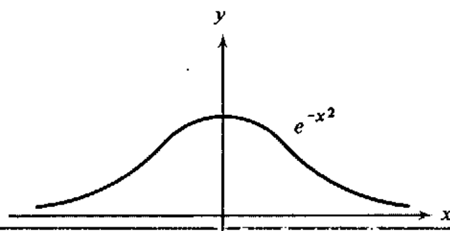


FIGURA 9.ER-1 La función gaussiana

Solución Para usar coordenadas polares, sea $I_b = \int_{A_b} e^{-x^2-y^2}\,dx\,dy$, donde A_b es el círculo de radio b con centro en el origen. Así,

$$I_b = \int_{r=0}^{b}\int_{\theta=0}^{2\pi} e^{-r^2}\,r\,d\theta\,dr,$$

lo que evaluamos mediante integrales iteradas. Como $(d/dr)(e^{-r^2}) = -2re^{-r^2}$, obtenemos

$$I_b = 2\pi\int_{r=0}^{b} re^{-r^2}\,dr = \pi(1 - e^{-b^2}).$$

Para relacionar I_b con $\int_{-\infty}^{\infty} e^{-x^2}\,dx$, primero observamos que

$$\int_{\mathbb{R}^2} e^{-x^2-y^2}\,dx\,dy = \lim_{b\to\infty} \pi(1 - e^{-b^2}) = \pi.$$

Como la integral existe y $e^{-x^2-y^2} \geq 0$, podemos evaluarla de la forma que queramos. Vamos a evaluar $\int_{\mathbb{R}^2} e^{-x^2-y^2}\,dx\,dy$ mediante el rectángulo $[-b, b]^2 = [-b, b] \times [-b, b]$. Así,

$$\int_{\mathbb{R}^2} e^{-x^2-y^2}\,dx\,dy = \lim_{b\to\infty}\int_{[-b,b]^2} e^{-x^2-y^2}\,dx\,dy.$$

Pero

$$\int_{[-b,b]^2} e^{-x^2-y^2}\,dx\,dy = \left(\int_{-b}^{b} e^{-x^2}\,dx\right)\left(\int_{-b}^{b} e^{-y^2}\,dy\right) = \left(\int_{-b}^{b} e^{-x^2}\,dx\right)^2$$

por el teorema de Fubini y el hecho de que $e^{-x^2-y^2} = e^{-x^2}e^{-y^2}$. Concluimos que $\lim_{b\to\infty} (\int_{-b}^{b} e^{-x^2}\,dx)^2 = \pi$. Como $e^{-x^2} \geq 0$, tenemos que $\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$, como se pedía. ♦

Ejemplo 9.2 Regla de Leibniz para la derivación de integrales *Supóngase que $f: [a, b] \times [c, d] \to \mathbb{R}$ es continua y que $\partial f/\partial y$ existe y es continua en $[a, b] \times [c, d]$. Sean $u(t)$ y $v(t)$ funciones continuamente derivables de $[c, d]$ en $[a, b]$ y sea*

$$F(t) = \int_{u(t)}^{v(t)} f(x, t)\,dx.$$

Demuéstrese que F es derivable y que

$$F'(t) = f(v(t), t)v'(t) - f(u(t), t)u'(t) + \int_{u(t)}^{v(t)} \frac{\partial f}{\partial t}(x, t)\,dx.$$

Solución Ésta es una aplicación de la regla de la cadena, una vez planteada de forma adecuada. Sean $u = u(t)$, $v = v(t)$ y $w = t$. Entonces $F(t) = g(u, v, w) = \int_u^v f(x, w)dx$. Así,

$$F'(t) = \frac{\partial g}{\partial u}\frac{du}{dt} + \frac{\partial g}{\partial v}\frac{dv}{dt} + \frac{\partial g}{\partial w}\frac{dw}{dt}.$$

Las derivadas parciales de g con respecto a u y v se obtienen mediante el teorema fundamental del cálculo. Como ese teorema utiliza el límite superior de la integral, la derivación con respecto al límite inferior se realiza invirtiendo los límites y aplicando el teorema. Esta inversión introduce un signo menos. La derivada parcial con respecto a w se obtiene derivando bajo el signo de integral, mediante la proposición **9.7.4.** Como $w = t$ y $dw/dt = 1$, esto nos da la fórmula deseada. ♦

Ejemplo 9.3 *Calcúlese el volumen de la bola de radio r con centro en el origen en $\mathbb{R}^n$ (es decir, el conjunto $\{x + \mathbb{R}^n \mid \|x\| < r\}$).*

Solución Usamos inducción sobre la dimensión de n. En $\mathbb{R}$, la bola es simplemente el intervalo abierto $]-r, r[$ y tiene volumen $2r$. Supóngase que hemos calculado el volu-

men de la bola de radio r en $\mathbb{R}^{n-1}$ como $a_{n-1}\, r^{n-1}$ (se intuye que la respuesta es de esta forma, pues la bola en $\mathbb{R}^{n-1}$ es un objeto $(n-1)$-dimensional). Entonces, como la frontera de la bola en $\mathbb{R}^{n-1}$ tiene medida nula, podemos aplicar el teorema de Fubini. Para cada x_n fijo, $0 \le x_n < r$, la sección transversal de la bola de radio r en $\mathbb{R}^{n-1}$, que denotamos $B(n, r)$, es

$$\{(x_1, \ldots, x_{n-1}, x_n) \mid x_1^2 + \cdots + x_{n-1}^2 < r^2 - x_n^2\},$$

que es una bola de radio $(r^2 - x_n^2)^{1/2}$ en $\mathbb{R}^{n-1}$. Por lo tanto, el teorema de Fubini implica

$$\int_{B(n,r)} 1 = \int_{-r}^{r} \int_{B(n-1,(r^2-x_n^2)^{1/2})} 1 = \int_{-r}^{r} a_{n-1}((r^2 - x_n^2)^{1/2})^{n-1}\, dx_n.$$

Haciendo $x_n = r \operatorname{sen} \theta$ para $0 < \theta < \pi/2$, obtenemos $(r^2 - x_n^2)^{1/2} = r \cos \theta$ y $dx_n/d\theta = r \cos \theta > 0$ en $]0, \pi/2[$. Así,

$$\begin{aligned}
\int_{B(n,r)} 1 &= \int_{-r}^{r} a_{n-1}((r^2 - x_n^2)^{1/2})^{n-1}\, dx_n = 2 \int_{0}^{r} a_{n-1}((r^2 - x_n^2)^{1/2})^{n-1}\, dx_n \\
&= 2 \int_0^{\pi/2} a_{n-1}(r \cos \theta)^{n-1} r \cos \theta \, d\theta = 2a_{n-1} r^n \int_0^{\pi/2} \cos^n \theta \, d\theta \\
&= a_n r^n, \qquad \text{donde } a_n = 2a_{n-1} \int_0^{\pi/2} \cos^n \theta \, d\theta.
\end{aligned}$$

Usando cálculo elemental, tenemos que

$$\int_0^{\pi/2} \cos^n \theta \, d\theta = \begin{cases} \dfrac{(n-1)(n-3)\cdots}{n(n-2)\cdots} & n \text{ impar}, \\[2ex] \dfrac{(n-1)(n-3)\cdots}{n(n-2)\cdots} \cdot \dfrac{\pi}{2} & n \text{ par}. \end{cases}$$

Por lo tanto, tenemos

$$\begin{aligned}
v(B(1,r)) &= 2r & v(B(2,r)) &= 2 \cdot a_1 \cdot r^2 \cdot \frac{1}{2} \cdot \frac{\pi}{2} \\
& & &= \pi r^2 \\
v(B(3,r)) &= 2 \cdot a_2 \cdot r^3 \cdot \frac{2}{3} & v(B(4,r)) &= 2 \cdot a_3 \cdot r^4 \cdot \frac{3}{8} \cdot \frac{\pi}{2} \\
&= \frac{4}{3} \pi r^3 & &= \frac{\pi^2}{2} r^4 \\
v(B(5,r)) &= 2 \cdot a_4 \cdot r^5 \cdot \frac{8}{15} & v(B(6,r)) &= 2 \cdot a_5 \cdot r^6 \cdot \frac{15}{48} \cdot \frac{\pi}{2} \\
&= \frac{8}{15} \pi^2 r^5 & &= \frac{1}{6} \pi^3 r^6. \quad \blacklozenge
\end{aligned}$$

Ejemplo 9.4 *Mejórese la fórmula del cambio de variables reemplazando "A es abierto" por "A tiene volumen".*

Solución Existen (al menos) dos formas de hacer esto, y daremos ambas como teoremas.

Teorema *Sea $g : D \subset \mathbb{R}^n \to \mathbb{R}^n$ de clase C^1, donde D es abierto. Además, sea G inyectiva y $Jg(x) \neq 0$ para todo $x \in D$. Sea $B = g(D)$. Supóngase que D y B tienen volumen. Sea $A \subset B$ con volumen y $f: A \to \mathbb{R}$ integrable. Entonces*

$$\int_{g^{-1}(A)} (f \circ g)|Jg| = \int_A f.$$

Demostración Extendemos f a B haciendo $f = 0$ fuera de A. Por **9.3.1**, $\int_D (f \circ g) \, |Jg| = \int_B f$. Como $f = 0$ fuera de A, $f \circ g$ se anula fuera de $g^{-1}(A)$, de donde se sigue la conclusión. ■

Teorema *Sea B con volumen y $f: B \to \mathbb{R}$ integrable. Sea A con volumen y supóngase que g : $\mathrm{int}(A) \subset \mathbb{R}^n \to \mathrm{int}(B) \subset \mathbb{R}^n$ es C^1 y biyectiva, y que $Jg(x) \neq 0$ para todo $x \in$ $\mathrm{int}(A)$. Si $f: B \to \mathbb{R}$ es integrable, entonces*

$$\int_B f = \int_A (f \circ g)|Jg|.$$

Demostración Como B tiene volumen y $\partial(\mathrm{int}(B)) \subset \partial(B)$, $\mathrm{int}(B)$ tiene volumen (pues $\partial(B)$ tiene medida nula). Además, $\mathrm{int}(B) \cup (\partial(B) \cap B) = B$, de modo que $\int_{\mathrm{int}(B)} f = \int_B f$. Por lo tanto, obtenemos el resultado por el teorema del cambio de variables. ■

Obsérvese que las condiciones sobre g son equivalentes a la existencia de una inversa C^1 para g (por el teorema de la función inversa). ♦

Nota. En estos dos teoremas, se puede mostrar que la hipótesis $Jg(x) \neq 0$ se puede eliminar (entonces, g no tendrá necesariamente una inversa C^1). Esto se puede deducir del ejercicio 27. El ejercicio 26 pide al lector que demuestre la fórmula del cambio de variables para integrales impropias. Una solución se puede basar en la fórmula del cambio de variables usual y nuestro análisis de las integrales impropias en el capítulo 8.

Ejercicios del capítulo 9

1. Úsense las coordenadas cilíndricas $g(r, \theta, z) = (r\cos\theta, r\,\text{sen}\,\theta, z)$ definidas en $\{r, \theta, z) \mid r > 0, 0 < \theta < 2\pi\}$ para calcular la integral sobre $A = \{(x, y, z) \mid x^2 + y^2 < 1, |z| < 1\}$ de $f(x, y, z) = (x^2 + y^2)z^2$.

2. Proporciónese un contraejemplo para mostrar que la fórmula del cambio de variables no es válida si g no es inyectiva, aunque $Jg(x) \neq 0$. [Sugerencia: sea $f = 1$ y $g(x, y) = (e^x \cos y, e^x \,\text{sen}\, y)$.]

3. Evalúense las siguientes integrales, si existen:

a. $\int_A x^2y^2\,dx\,dy$, donde $A = \{(x,y) \mid 0 < x < y^2, 0 < y < 2+x, x < 1\}$.

b. $\int_A \text{sen}(x^2 + y^2)\,dx\,dy$, donde A es el disco unidad.

c. $\int_{\mathbb{R}^3} 1/(x^2 + y^2 + z^2)\,dx\,dy\,dz$.

d. $\int_A (y/\sqrt{x})\,dx\,dy$, donde A es el cuadrado unidad $\{x, y) \mid 0 < x < 1, 0 < y < 1\}$.

e. $\int_A x\,dx\,dy$, donde $A = \{(x,y) \mid 0 < x < \sqrt{\pi}, 0 < y < \text{sen}\,x^2\}$.

f. $\int_0^\pi \int_0^1 r^2\,dr\,d\theta$.

g. $\int_{-1}^1 \int_0^{(1-x^2)^{1/2}} (x^3 + 3y^2x)\,dy\,dx$.

4. Supóngase que $f, f_1, f_2, f_3, \ldots$ son funciones continuas de valores reales en $[0, 1]$ y que $f_n \to f$ uniformemente en $[0, 1]$ cuando $n \to \infty$. Demuéstrese que cada una de las funciones $|f|, |f_1|, |f_2|, |f_3|, \ldots$ es integrable en $[0, 1]$ y que $\int_0^1 |f_n| \to \int_0^1 |f|$ cuando $n \to \infty$.

5. Calcúlese el volumen de los siguientes conjuntos:

a. Un tetraedro con área de la base A y altura h.

b. Un cono con radio de la base r_0 y altura h_0.

c. $\{(x,y) \mid x^2 < y < 1 - x^2\}$.

d. $\{(x,y,z) \mid x^2 + y^2 + z^2 < 1 \;\text{ y }\; z < 1/2\}$.

6. Demuéstrese:

a. Si A tiene volumen y si λ se define como

$$\lambda = \inf\left\{\sum_{i=1}^{\infty} v(S_i) \mid S_1, S_2, \ldots \text{ es un recubrimiento numerable de } A \text{ mediante rectángulos abiertos}\right\}, \text{ entonces } v(A) = \lambda,$$

b. Sea A un conjunto acotado con volumen y A_i una sucesión de conjuntos con volumen tales que los A_i no se solapan (es decir, sus interiores no se intersecan) y

$$\bigcup_{i=1}^{\infty} A_i = A.$$

Muéstrese que

$$v(A) = \sum_{i=1}^{\infty} v(A_i).$$

7. Determínese $\int_A xy \operatorname{sen}(x^2 - y^2)\, dx\, dy$, donde

$$A = \{(x, y) \mid 0 < y < 1, x > y, \text{ y } x^2 - y^2 < 1\}.$$

8. **a.** Sean $u : A \subset \mathbb{R}^2 \to]a, b[$ y $v : B \subset \mathbb{R}^2 \to]c, d[$ dos funciones de clase C^1 de los conjuntos abiertos A y B sobre los intervalos $]a, b[$ y $]c, d[$ tales que $u(x, y) = u(x', y')$ y $v(x, y) = v(x', y')$ solamente cuando $(x, y) = (x', y')$ y que

$$\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial v}{\partial x}\frac{\partial u}{\partial y} \neq 0$$

en cualquier punto $(x, y) \in A \cap B$. Sea $W = \{(x, y) \mid a < u(x, y) < b, c < v(x, y) < d\}$ y sea f una función integrable en W. Muéstrese que

$$\int_W f = \int_c^d \int_a^b f(u, v) \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial v}{\partial x}\frac{\partial u}{\partial y}\right)^{-1} du\, dv.$$

b. Úsese **a** para evaluar

$$\int_W (x^2 + y^2)\, dx\, dy,$$

donde $W = \{(x, y) \mid x > 0, y > 0, -1 < x^2 - y^2 < 1, xy < 1\}$.

9. Supóngase que $f :]a, b[\to \mathbb{R}$ y $g :]c, d| \to \mathbb{R}$ son dos funciones integrables. Defínase $\tilde{f}(x, y) = f(x)$, $\tilde{g}(x, y) = g(y)$ (supóngase que f y g están acotadas). Demuéstrese que

$$\int_{[a,b]\times[c,d]} \tilde{f}(x, y)\tilde{g}(x, y)\, dx\, dy = \left(\int_a^b f(x)\, dx\right)\left(\int_c^d g(x)\, dx\right)$$

10. Supóngase que $A \subset \mathbb{R}^n$ y que A tiene volumen nulo. Supóngase que $f : A \to \mathbb{R}$ es una función acotada. Demuéstrese que f es integrable y que $\int_A f = 0$.

11. Sea $S = \{(x, y) \in \mathbb{R}^2 \mid x$ racional, $0 < x < 1$ y si $x = p/m$ fracción irreducible, $y = k/m$, $k = 1, \ldots, m-1\}$. Muéstrese que

$$\int_0^1 \int_0^1 1_S \, dy \, dx = 0$$

pero que

$$\int_{[0,1]\times[0,1]} 1_S$$

no existe.

12. Muéstrese que si $f''(x) > 0$ para todo x, entonces f es *convexa hacia arriba, es decir,*

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2}$$

13. Supóngase que $C \subset A \times B$ y $v(C) = 0$. Sea $C_x = \{y \in B \mid (x, y) \in C\}$ y suponga que

$$1_{C_x}(y) = \begin{cases} 1 & \text{si } (x,y) \in C \\ 0 & \text{en caso contrario} \end{cases}$$

es integrable sobre B para cada $x \in A$. Muéstrese que C_x tiene volumen nulo en A para todo x excepto quizá un conjunto de medida nula. Proporciónese un ejemplo donde $v(C_x) \neq 0$ para algún x.

14. *Verdadero o falso* (si es verdadero, dése una razón; si es falso, formúlese un contraejemplo).

a. Supóngase que f es integrable en A y que g es una función en A tal que $g \leq f$. Entonces g también es integrable en A.

b. Supóngase que A tiene volumen y que f es continua en A. Entonces $\int_A f$ existe.

c. Supóngase que A tiene volumen y que $\int_A f$ existe. Entonces f es continua en A.

d. Cualquier subconjunto acotado de $\mathbb{R}^2$ tiene volumen nulo cuando se considera como un subconjunto de $\mathbb{R}^3$.

e. Si $I = [0, 1]$ y φ_1 y φ_2 son funciones continuas de I en $\mathbb{R}$ con $\varphi_1(x) < \varphi_2(x)$ para todo $x \in I$, entonces $S = \{(x, y) \mid x \in I, \varphi_1(x) < y < \varphi_2(x)\}$ tiene volumen.

15. Si A es un conjunto acotado con volumen y A_i es una sucesión de conjuntos con volumen tal que $A_{i+1} \supset A_i$ y $A_1 \cup A_2 \cup \cdots = A$, muéstrese entonces que $v(A_i) \to v(A)$ cuando $i \to \infty$.

16. Supóngase que $g: \mathbb{R} \to \mathbb{R}$ es derivable en todo punto y que $|g'(x)| \leq M$ para todo $x \in \mathbb{R}$. Muéstrese que si ε es suficientemente pequeño, la función f dada por $f(x) = x + \varepsilon g(x)$ es inyectiva.

17. Sea $f: [0, 1] \to \mathbb{R}$ dada por

$$f(x) = \begin{cases} n^2 & \text{si } x = 1/n \\ 0 & \text{si } x \in [0, 1],\ x \neq 1/n. \end{cases}$$

Demuéstrese que f es integrable y que $\int_0^1 f(x)\, dx = 0$. (*Advertencia:* f no está acotada.)

18. Sea $f_n(x) = \sum_{m=1}^{n} (1/2^m)$ sen mx definida para todo $x \in \mathbb{R}$.

a. Muéstrese que $\lim_{n\to\infty} f_n(x)$ existe.

b. Muéstrese que la sucesión converge uniformemente.

c. Muéstrese que $\int_0^{2\pi} (\lim_{n\to\infty} f_n(t))\, dt = 0$.

19. Sean $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^m$ con volumen y $f: A \to \mathbb{R}$, $g: B \to \mathbb{R}$ integrables. Sea $F(x, y) = f(x) + g(y)$. Muéstrese que

$$\int_{A\times B} F(x, y)\, dx\, dy = \left(\int_A f\right) v(B) + \left(\int_B g\right) v(A).$$

20. Calcúlese

a. $\lim_{n\to\infty} (1 + 2^k + \cdots + n^k)/n^{k+1}$, donde $k > 0$.

b. $\lim_{n\to\infty} \left(1/(n+1) + 1/(n+2) + \cdots + 1/2n\right)$.

21. ¿Para qué valores de p es r^p integrable en $\mathbb{R}^3$, donde

$$r = \sqrt{x^2 + y^2 + z^2}?$$

22. Sean f y g funciones de valores reales, integrables en $[a, b]$. Sea

$$h(x) = \text{máx}\ (f(x), g(x)) = \begin{cases} f(x) & \text{si } f(x) \geq g(x), \\ g(x) & \text{si } f(x) \leq g(x). \end{cases}$$

Muéstrese que $h(x)$ es integrable.

23. Sea A un rectángulo cerrado en $\mathbb{R}^n$ y $C \subset A$ con volumen. Demuéstrese que para cualquier $\varepsilon > 0$ existe un conjunto compacto $K \subset C$ tal que $v(C \backslash K) < \varepsilon$, y que existe un conjunto compacto $L \supset C$ tal que $v(L \backslash C) < \varepsilon$.

24. *Verdadero o falso:*

a. Si f es una función continua en $[0, 1]$, entonces f está acotada en $[0, 1]$.

b. Si $f: S \to \mathbb{R}^n$ es una función continua, donde S es un subconjunto compacto de $\mathbb{R}$, entonces $f(S)$ es compacto.

c. Una función integrable en $[0, 1]$ debe ser continua en $[0, 1]$.

d. Si U y V son subconjuntos abiertos de $\mathbb{R}$, entonces $U \times V = \{(x, y) \mid x \in U, y \in V\}$ es abierto en $\mathbb{R}^2$.

e. Si f y g son funciones integrables en $[a, b]$, entonces $f - 2g$ es integrable en $[a, b]$.

f. Cualquier sucesión acotada en $\mathbb{R}^n$ debe tener una subsucesión convergente.

g. Si f es una función continua de valores reales en $[0, 1]$ tal que $f(x) \geq 0$ para todo $x \in [0, 1]$ y $\int_0^1 f(x)\, dx = 0$, entonces $f(x) = 0$ para todo $x \in [0, 1]$.

h. Si f es una función de valores reales infinitamente derivable en $\mathbb{R}$, entonces f tiene un desarrollo en serie de potencias en torno de cada punto de $\mathbb{R}$.

i. Si $f, f_1, f_2, f_3, \ldots$ son funciones de valores reales, continuas y uniformemente acotadas en $[0, 1]$ tales que $f(x) = \lim_{n\to\infty} f_n(x)$ para cada $x \in [0, 1]$, entonces $f_1, f_2, f_3, \ldots$ converge a f uniformemente en $[0, 1]$.

j. Si $a_0, a_1, a_2, \ldots$ es una sucesión de números reales y r es el radio de convergencia de $\sum_{n=0}^{\infty} a_n x^n$, entonces $\sum_{n=0}^{\infty} a_n r^n$ converge.

k. Si $f_1, f_2, f_3, \ldots$ converge uniformemente a f en $[a, b]$ y si f_n es integrable para cada $n = 1, 2, 3, \ldots$, entonces f es integrable en $[a, b]$.

l. Si $f_1, f_2, f_3, \ldots$ converge uniformemente a f en $[a, b]$ y si f_n es derivable en $]a, b[$ para cada $n = 1, 2, 3, \ldots$, entonces f es derivable en $[a, b]$.

m. Un subconjunto abierto conexo de $\mathbb{R}^n$ es conexo por arcos.

n. Si f es una función derivable de valores reales en $]0, 1[$ y $f(1/2) \geq f(x)$ para todo $x \in]0, 1[$, entonces $f'(1/2) = 0$.

o. Si f es una función integrable en $[0, 1]$ y $\varepsilon > 0$, entonces existe una función escalonada g en $[0, 1]$ tal que $\int_0^1 |f - g| < \varepsilon$.

p. Todo subconjunto cerrado y acotado de $\mathbb{R}$ contiene a su supremo.

q. Si $f(x)$ tiene un desarrollo en serie de potencias en $]-r, r[$, entonces f es derivable en $]-r, r[$.

r. Si f es integrable en $[a, b]$ y $[b, c]$, donde $a < b < c$, entonces f es integrable en $[a, c]$.

s. Si $f : \mathbb{R}^n \to \mathbb{R}^m$ es continua y S es cerrado en $\mathbb{R}^m$, entonces $f^{-1}(S)$ es cerrado en $\mathbb{R}^n$.

t. Si una serie $\sum_{j=1}^{\infty} a_j$ converge condicionalmente, entonces la sucesión de sumas parciales $\sum_{j=1}^{n} a_j$ está acotada.

u. Si f es una función continua de valores reales en $]a, b[$, entonces existe una función derivable F en $]a, b[$ tal que $f = F'$ en $]a, b[$.

25. Calcúlese el área de la región $D = \{(x, y) \mid 1 \leq x < 3$ y $x^2 \leq y \leq x^2 + 1$.

26. Investíguense las posibles generalizaciones del teorema **9.3.1** a regiones no acotadas.

27. El propósito de este problema es demostrar una versión simplificada de un teorema algo difícil, conocido como el ***teorema de Sard***. Un tratamiento general aparece en Milnor, *Topology from the Differentiable Viewpoint*, Princeton University Press (1967), o en Abraham, Marsden y Ratiu, *Manifolds, Tensor Analysis and Applications*, Springer-Verlag (1988). En nuestro caso, el enunciado del teorema es el siguiente.

Teorema *Sea $A \subset \mathbb{R}^n$ un conjunto abierto y $g : A \to \mathbb{R}^n$ de clase C^1. Sea $B = \{x \in A \mid Jg(x) = 0\}$. Entonces $g(B)$ tiene medida nula.*

El conjunto B no tiene por qué ser de medida nula (para verlo, tómese g como una función constante). Antes de dar un esquema de la demostración, pediremos al lector que dé por válido el resultado y lo use para mostrar que en el teorema **9.3.1** se puede eliminar la hipótesis de que $Jg(x) \neq 0$ (siempre que el conjunto (abierto) de puntos tales que $Jg(x) \neq 0$ tenga volumen).

El teorema se demuestra como sigue. Primero se muestra que si U es un rectángulo cerrado en A, basta mostrar que $g(U \cap B)$ tiene medida nula (muéstrese que $g(B)$ es la unión numerable de estas intersecciones). Mostraremos, de hecho, que $g(U \cap B)$ tiene contenido nulo.

Demuéstrense estos dos puntos: *para cualquier $\varepsilon > 0$, existe $\delta > 0$ tal que para x, $y \in U$, $\|x - y\| < \delta$, tenemos*

$$\|g(x) - g(y) - \mathbf{D}g(x)(x - y)\| < \varepsilon\|x - y\|.$$

Además, existe M tal que

$$\|g(x) - g(y)\| \leq M\|y - x\|.$$

Supongamos que U tiene lados de longitud l. Hágase $N \geq l/\delta$ tal que si U se divide en N^n rectángulos (de lados l/N) y S es uno de tales rectángulos, entonces, para $x, y \in S$ son válidas las dos desigualdades anteriores. Supóngase que $x \in S \cap B$. Determínese un hiperplano H en $\mathbb{R}^n$ (H será algún subespacio $(n-1)$-dimensional) tal que

$$\{\mathbf{D}g(x)(y - x) \mid y \in S\} \subset H.$$

A continuación, muéstrese que $\{g(y) \mid y \in S\}$ está contenido en un cilindro de altura $< 2\varepsilon n(l/N)$ y base un cubo $(n-1)$-dimensional de lado $< 2Mn(l/N)$. Dedúzcase que $g(U \cap B)$ está contenido en N^n rectángulos de volumen total $< \varepsilon K$, donde $K = 2^n M^{n-1} (n)^n l^n$ es una constante independiente de N. Esto demostrará el resultado.

28. Supóngase que $f(x)$ es continua en $]-1, 1[$, $f(0) = 0$ y $f(x) \neq 0$ si $x \neq 0$. Demuéstrese que $\lim_{x\to\infty} [\log(1 + f(x))]/f(x)$ existe. ¿A qué es igual?

29. Supóngase que f está acotada y definida en $[0, b]$, $b > 0$, y que $\int_c^b f$ existe para todo $0 < c < b$. Muéstrese que $\int_0^b f$ existe.

30. Considérese el siguiente teorema.

Teorema *Sea A un subconjunto abierto de $\mathbb{R}^n$ y $\varphi : A \to \mathbb{R}^n$ una transformación inyectiva, continuamente diferenciable, cuyo jacobiano $J\varphi$ no se anula en A. Supóngase que la función $f : \varphi(A) \to \mathbb{R}$ es continua, que se anula fuera de un subconjunto compacto de $\varphi(A)$ y que $\int_{\varphi(A)} f$ existe. Entonces $\int_{\varphi(A)} f = \int_A (f \circ \varphi) |J\varphi|$.*

Sean A_i, $i = 1, 2, \ldots$, subconjuntos abiertos de $\mathbb{R}^n$ tales que el teorema es cierto para cada A_i y para la restricción de φ a A_i. Sea $A = \cup A_i$. Muéstrese que el teorema es cierto para A y φ.

31. Calcúlese $\lim_{n\to\infty} n \log(1 + 1/n)$.

32. Supóngase que $S \subset \mathbb{R}^n$ tiene volumen y sea $t \in \mathbb{R}$. Sea $tS = \{tx_1, \ldots, tx_n) \mid (x_1, \ldots, x_n) \in S$. Muéstrese que tS tiene volumen y que $\mathrm{vol}(tS) = |t|^n \mathrm{vol}(S)$.

33. **a.** Demuéstrese el teorema **9.2.3.** [Sugerencia: como en el teorema **9.2.1**, basta demostrar **ii.** Hágase de la misma forma que en el teorema **9.2.1.**]

b. Demuéstrese la siguiente generalización del corolario **9.2.2**. Sea $A \subset \mathbb{R}^n$ un rectángulo cerrado y sean $\varphi : A \subset \mathbb{R}^n \to \mathbb{R}^m$ y $\psi : A \subset \mathbb{R}^n \to \mathbb{R}^m$ funciones

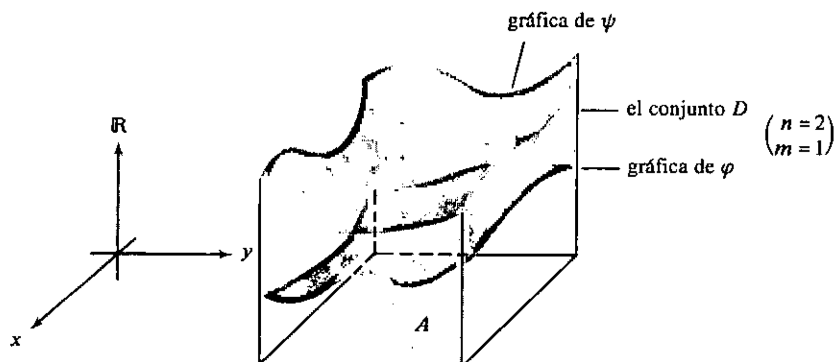


FIGURA 9.E-1 El conjunto D del ejercicio 33 es la región entre las gráficas de las funciones

continuas, tales que $\varphi_j(x) \leq \psi_j(x)$ para todo $x \in A$, $1 \leq j \leq m$. Sea $D = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid x \in A,\ \varphi_j(x) \leq y_j \leq \psi_j(x),\ 1 \leq j \leq m\}$. Véase la figura 9.E-1. Sea $f : A \to \mathbb{R}$ continua; defínase $B_x \subset \mathbb{R}^m$ como $B_x = \{y \in \mathbb{R}^m \mid \varphi_j(x) \leq y_j \leq \psi_j(x),\ 1 \leq j \leq m\}$, defínase $f_x : B_x \subset \mathbb{R}^m \to \mathbb{R}$ como $f_x(y) = f(x, y)$, y defínase $g : A \subset \mathbb{R}^n \to \mathbb{R}$ como $g(x) = \int_B f_x$. Entonces g es integrable en A y $\int_A g = \int_D f$.

Capítulo 10
Análisis de Fourier

El análisis de Fourier surgió históricamente en conexión con problemas de mecánica, como la conducción del calor y el movimiento ondulatorio. El tema se ha desarrollado como una amplia teoría con muchas aplicaciones, tanto matemáticas como físicas. La idea de este capítulo es dar un breve, pero básico, conocimiento operativo de algunos métodos de Fourier, introducir al estudiante en la teoría general y delinear algunas de sus aplicaciones fundamentales.

Como introducción al estudio del análisis de Fourier, consideremos una cuerda vibrante. En secciones posteriores se presentan más aplicaciones (por ejemplo, a la mecánica cuántica). Considérese una cuerda de longitud l, con extremos fijos, que se deja vibrar libremente después de pulsarse. La posición (desplazamiento vertical) de la cuerda se representa mediante una función $y(t, x)$, donde t es el tiempo y $x \in [0, l]$. Véase la figura 10-1. De la mecánica elemental se sabe que, para vibraciones de pequeña amplitud, y obedece la ***ecuación de ondas***

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2},$$

donde c es una constante determinada por la naturaleza de la cuerda y su tensión. El hecho de que la cuerda tenga fijos los extremos indica que $y(t, 0) = y(t, l) = 0$ para todo t.

Para simplificar nuestra exposición, analizaremos primero el caso de las soluciones particulares llamadas ***ondas estacionarias***; éstas son soluciones de la forma $y(t, x) = u(x)\cos \omega t$, donde $y(t, x)$, como en la figura 10-1, representa el desplazamiento vertical en x en el instante t, y donde ω es la frecuencia. Así, $|u(x)|$ representa la ***amplitud*** en el punto x, y la ***forma de la onda*** está dada por la función $u(x)$. Físicamente, una onda estacionaria es un movimiento síncrono hacia arriba y hacia abajo que repite su forma periódicamente después de un tiempo $t = 2\pi/\omega$, como el que ocurre cuando una cuerda produce una nota pura. Ondas estacionarias específicas, llamadas ***soluciones fundamentales*** o ***armónicos***, están dadas por

$$y_n(t, x) = \operatorname{sen}\left(\frac{n\pi x}{l}\right)\cos(\omega_n t), \quad n = 0, 1, 2, \ldots,$$

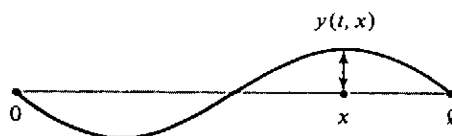


FIGURA 10-1 El desplazamiento de una cuerda en la posición x y en el instante t se denota $y(x, t)$

donde $\omega_n = n\pi c/l$ es la frecuencia. Por ejemplo, si $n = 2$ y $t = \pi/\omega_n$, obtenemos la onda de la figura 10-1.

Es importante y notable que cualquier solución $y(x, t)$ que describa el movimiento de la cuerda se puede descomponer en armónicos; es decir, se puede escribir como una serie

$$y(x,t) = \sum_{n=1}^{\infty} c_n y_n(x,t) = \sum_{n=1}^{\infty} c_n u_n(x) \cos(\omega_n, t),$$

donde y_n es como antes y $u_n(x) = \text{sen}(n\pi x/l)$. Decimos que u_1 es la ***primera componente armónica*** (o ***primer armónico***) de y, u_2 la ***segunda***, etcétera. Así, una vibración de apariencia complicada (como la que ocurre en las cuerdas de un violín) es en realidad una combinacion infinita de armónicos, donde cada componente armónica aparece con peso c_n. En la figura 10-2 mostramos cómo la suma de tres curvas sinusoidales con amplitudes variables puede dar lugar a una forma de onda más complicada. Para conseguir una forma de onda general se podría necesitar una combinación infinita de curvas sinusoidales.

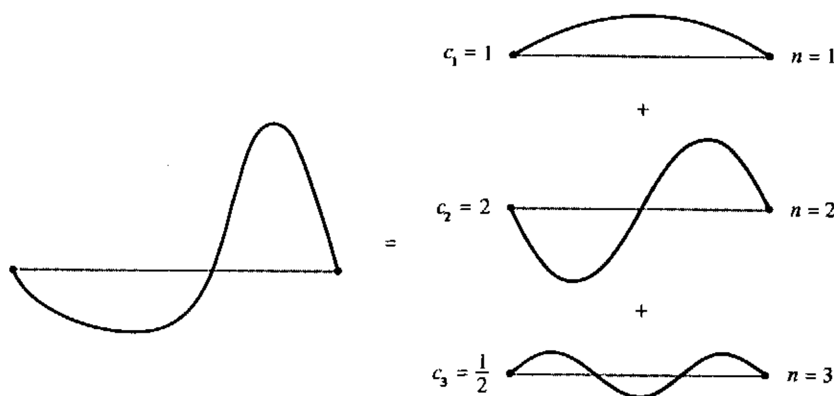


FIGURA 10-2 La superposición de tres armónicos

La finalidad del análisis de Fourier es realizar este procedimiento de descomposición mediante un método general. Para regiones finitas (como el intervalo $[0, l]$ del

ejemplo anterior), el método adecuado es el uso de *series de Fourier*, mientras que en una región infinita (toda la recta real, por ejemplo), se necesitan *integrales de Fourier*.

Las series obtenidas para sen nx, cos nx o e^{inx} son las ***series de Fourier clásicas***. Para otros tipos de problemas (el oscilador armónico en la mecánica cuántica, por ejemplo), aparecen otros tipos de soluciones básicas, y cualquier solución arbitraria debe desarrollarse en términos de estas soluciones básicas (por ejemplo, para el oscilador armónico de la mecánica cuántica se usan las funciones de Hermite). En consecuencia, es útil analizar la teoría general de tales desarrollos, como lo haremos en **§10.1** y **§10.2**.

En **§10.3**, **§10.4** y **§10.5** estudiaremos el caso particular de las series de Fourier trigonométricas y justificaremos el procedimiento de desarrollo. Para justificar este procedimiento en el caso de otras familias de funciones, a menudo hay que examinar las situaciones particulares (por ejemplo, la ecuación diferencial que da origen al problema). A este respecto existen dos teoremas principales. El primero trata el importante concepto de convergencia en media y establece que cualquier función de cuadrado integrable tiene una serie de Fourier que converge en media. Como explicaremos más adelante, esto no debe confundirse con la convergencia puntual. Para esta última, podemos usar el teorema básico de Dirichlet y Jordan. En **§10.6** daremos otros teoremas de convergencia, como la justificación de la derivación término a término. En **§10.7** presentamos unas cuantas aplicaciones importantes de los métodos de Fourier. Ahí estudiamos algunos casos particulares de tres problemas: la ecuación de ondas, el problema de Dirichlet (ecuación de Laplace) y la ecuación del calor, todo desde el punto de vista de las series de Fourier. En **§10.8** damos un tratamiento informal de las integrales de Fourier, estableciendo sus propiedades y definiciones básicas, sin demostraciones (que se dejan para un curso más avanzado). Esperamos que el estudiante obtenga con esto cierta perspectiva del papel de las integrales de Fourier y su relación con las series de Fourier. Por último, en **§10.9** damos un pequeño vistazo a la mecánica cuántica y a la forma de usar las técnicas de **§10.1** a **§10.3** para establecer algunos de los resultados básicos de esta teoría.

§10.1 Espacios con producto interno

Antes de iniciar nuestro estudio de las series de Fourier analizaremos algunos conceptos que nos permitirán simplificar nuestro trabajo. Estas ideas revelan un importante aspecto geométrico de las series de Fourier. En este capítulo también necesitaremos algunos hechos básicos acerca de los números complejos del capítulo 1.

En el capítulo 1 estudiamos el producto interno $\langle , \rangle$ en $\mathbb{R}^n$. Ahora ampliaremos estos conceptos a un espacio vectorial arbitrario V. En este capítulo, V ya no será de dimensión finita, sino un espacio de funciones de dimensión infinita, como el espacio $C(A, \mathbb{R}^m)$ estudiado en el capítulo 5. Por ejemplo, V podría ser un espacio de funciones $f : [0, 2\pi] \to \mathbb{R}$. Éste es un espacio vectorial si usamos las definiciones usuales, $(f + g)(x) = f(x) + g(x)$ y $(\alpha f)(x) = \alpha f(x)$. La expresión $\langle f, g\rangle = \int_0^{2\pi} f(x)g(x)dx$ es el ***escalar***

(o ***producto interno***) de f y g. Éste es el tipo de espacio V que debe tener presente el lector al estudiar las siguientes dos secciones.

Es importante permitir que los elementos de nuestro espacio de funciones tomen valores complejos, por la sencilla razón de que, con frecuencia, es más conveniente trabajar con $e^{i\theta}$ que con $\operatorname{sen}\theta$ y $\cos\theta$. En el caso complejo, definimos

$$\langle f, g\rangle = \int_0^{2\pi} f(x)\overline{g(x)}\,dx,$$

donde $\overline{g(x)}$ es el complejo conjugado de $g(x)$. La razón por la que usamos $\overline{g(x)}$ es para que podamos (como en el caso real) definir la ***longitud*** o ***norma*** de f como

$$\|f\|^2 = \langle f,f\rangle = \int_0^{2\pi} |f(x)|^2\,dx$$

(para los números complejos z, recuérdese que $|z|^2 = z\bar{z}$ es un número real positivo).

Empezaremos nuestro estudio utilizando espacios vectoriales abstractos con producto interno, en vez del espacio de funciones particular ya mencionado (que es en realidad el más importante para nosotros), puesto que, en términos de conceptos y notaciones, es más sencillo trabajar con la notación $\langle\ ,\ \rangle$ que con integrales. En este momento sólo nos interesan las siguientes propiedades básicas de $\langle\ ,\ \rangle$.

10.1.1 Definición *Sea V un espacio vectorial complejo (es decir, un espacio vectorial en el que permitimos el uso de números complejos para la multiplicación por un escalar). Un **producto escalar** (o **interno**) en V es una aplicación $\langle\ ,\ \rangle : V \times V \to \mathbb{C}$ (donde $\mathbb{C}$ denota los números complejos), con las siguientes propiedades:*

i. $\langle \alpha f + \beta g, h\rangle = \alpha\langle f, h\rangle + \beta\langle g, h\rangle$ *para todo* $f, g, h \in V$ *y* $\alpha, \beta \in \mathbb{C}$.

ii. $\langle f, h\rangle = \overline{\langle h, f\rangle}$.

iii. $\langle f, f\rangle \geq 0$, *y* $\langle f, f\rangle = 0$ *implica que* $f = 0$.

De **i** y **ii** deducimos que $\langle h, \alpha f + \beta g\rangle = \overline{\alpha}\langle h, f\rangle + \overline{\beta}\langle h, g\rangle$. Obsérvese que si todas las cantidades fueran reales, tendríamos las mismas propiedades del producto escalar usual de $\mathbb{R}^n$. En general, como ya hemos señalado, V no será de dimensión finita, por lo que no podremos usar ni bases ni matrices finitas.

10.1.2 Teorema *El espacio V de las funciones continuas $f : [a, b] \to \mathbb{C}$ es un espacio con producto interno si definimos*

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\,dx.$$

La integral de una función de valores complejos está definida como

$$\int_a^b f(x)\,dx = \int_a^b f_1(x)\,dx + i\int_a^b f_2(x)\,dx,$$

donde $f = f_1 + if_2$. Las propiedades de las integrales complejas son similares a las de las integrales reales y se pueden deducir de éstas.

En un espacio con producto interno podemos introducir muchas de las ideas tratadas en el capítulo 1. La ***norma*** de f, que se denota $\|f\|$, se define como

$$\|f\|^2 = \langle f, f\rangle$$

y la ***distancia*** entre f y g como

$$d(f, g) = \|f - g\|.$$

Usamos el mismo lenguaje que en $\mathbb{R}^n$ por analogía. Por ejemplo, decimos que f y g son ***ortogonales*** si $\langle f, g\rangle = 0$. Como V es un espacio vectorial, también podemos hablar de la independencia lineal y otras ideas relacionadas con esto, que ya vimos en $\mathbb{R}^n$. El siguiente teorema sigue desarrollando la analogía con $\mathbb{R}^n$.

10.1.3 Desigualdad de Cauchy-Schwarz *Sean f, g pertenecientes al espacio con producto interno V; entonces*

$$|\langle f, g\rangle| \le \|f\|\,\|g\|.$$

Además, V con la norma $\|f\|$ satisface los axiomas de un espacio normado, y con $d(f, g) = \|f - g\|$, los axiomas de un espacio métrico.

Como tenemos un espacio métrico, serán importantes aquí los conceptos de conjuntos abiertos y cerrados del capítulo 2. En este caso, el concepto principal es el de convergencia de una sucesión o de una serie, que recordamos a continuación.

10.1.4 Definición *Sea V un espacio con producto interno y f_n una sucesión en V. Decimos que f_n* ***converge a*** *f y escribimos $f_n \to f$ si $\|f_n - f\| \to 0$, es decir, si para todo número real $\varepsilon > 0$ existe N tal que $n \ge N$ implica $\|f_n - f\| < \varepsilon$. Análogamente, una serie $\sum_{n=1}^{\infty} g_n$* ***converge a*** *f si la sucesión de sumas parciales $s_n = \sum_{k=1}^{n} g_k$ converge a f.*

Si V está formado por funciones $f: [a, b] \to \mathbb{C}$, entonces la desigualdad de Cauchy-Schwarz se lee

$$\left(\int_a^b f(x)\overline{g(x)}\,dx\right)^2 \le \left(\int_a^b |f(x)|^2\,dx\right)\left(\int_a^b |g(x)|^2\,dx\right).$$

La desigualdad triangular (es decir, $\|f+g\| \leq \|f\| + \|g\|$) se convierte en la llamada ***desigualdad de Minkowski*** y se lee

$$\left\{\int_a^b |f(x)+g(x)|^2\,dx\right\}^{1/2} \leq \left\{\int_a^b |f(x)|^2\,dx\right\}^{1/2} + \left\{\int_a^b |g(x)|^2\,dx\right\}^{1/2}.$$

Con este producto interno en el espacio de funciones V, la convergencia se denomina ***convergencia en media***. No necesariamente coincide con la convergencia puntual o uniforme, por lo que escribimos $f_n \to f$ (en media), $f_n \to f$ (puntualmente) y $f_n \to f$ (uniformemente) para distinguirlas. Como vimos en el capítulo 5, la convergencia uniforme implica la convergencia en media; sin embargo, la convergencia puntual no implica necesariamente la convergencia en media. En **§10.3** estudiaremos más relaciones entre los diversos tipos de convergencia.

Así, $f_n \to f$ (en media) tiene el mismo significado que el enunciado

$$\left(\int_a^b |f_n(x) - f(x)|^2\,dx\right) \to 0.$$

Por ejemplo, considérese la función f_n definida como $f_n(x) = 1 - nx$ para $0 \leq x \leq 1/n$ y $f_n(x) = 0$ para los demás $x \in [0, 1]$. Entonces $f_n \to 0$ (en media), pero f_n no converge a $f = 0$ puntualmente (en $x = 0$ en particular), por lo que tampoco converge uniformemente. Véase la figura 10.1-1. Podríamos considerar otros tipos de convergencia, como $\int_a^b |f_n(x) - f(x)|\,dx \to 0$, pero la convergencia en media es la más adecuada para el análisis de Fourier debido a su estrecha relación con los espacios con producto interno.

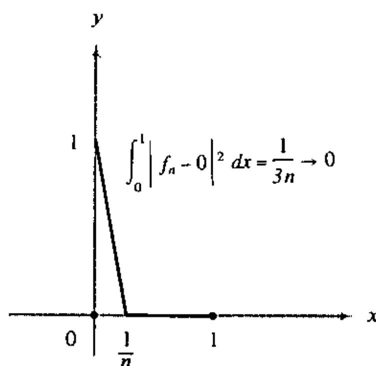


FIGURA 10.1-1 La sucesión f_n converge a cero en media, pero no puntualmente

El espacio V del teorema **10.1.2** se puede ampliar para incluir otras funciones, como las funciones continuas a trozos. Sin embargo, aunque V se extienda a todas las funciones integrables Riemann, adolece de una grave deficiencia: *no es completo*; es decir, existen sucesiones de Cauchy que no convergen. Recordemos este concepto.

10.1.5 Definición *Una sucesión f_n en un espacio con producto interno V es una* ***sucesión de Cauchy*** *si para todo $\varepsilon > 0$ existe N tal que $m,n \geq N$ implica $\|f_n - f_m\| < \varepsilon$. Un espacio con producto interno es* ***completo*** *si toda sucesión de Cauchy en dicho espacio converge. Un espacio con producto interno que sea completo es un* ***espacio de Hilbert***.

Para que V en el teorema **10.1.2** sea completo se usa la integral de Lebesgue. Nuestro análisis elemental no requiere de este concepto, pero el lector debe ser consciente de que hay una solución a este problema.

¿Podemos trabajar con funciones no continuas y aun así tener un espacio con producto interno? La respuesta es muy simple. El único punto del teorema **10.1.2** donde se usa la continuidad es en la afirmación

$$\int_a^b |f(x)|^2\,dx = 0 \qquad \text{implica} \qquad f = 0.$$

Para una función integrable f, hemos visto en el capítulo 8 que

$$\int_a^b |f(x)|^2\,dx = 0 \qquad \text{implica} \qquad f = 0,$$

excepto posiblemente para x perteneciente a un conjunto de medida nula. Si consideramos una tal f equivalente a cero (redefinimos f en un conjunto de medida nula en caso necesario), entonces el teorema **10.1.2** se puede aplicar. Con esta convención, es válido el siguiente teorema.

10.1.6 Teorema *Sea $V = L^2$ el espacio de funciones $f\colon [a, b] \to \mathbb{C}$ tales que $|f|^2$ es integrable (es decir, $\int_a^b |f(x)|^2 dx < \infty$). Entonces el espacio V es un espacio con producto interno, con el producto interno dado por*

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\,dx$$

y la norma

$$\|f\| = \left(\int_a^b |f(x)|^2\,dx\right)^{1/2}.$$

Otra clase conveniente de funciones que forman un espacio con producto interno es la clase las funciones continuas a trozos, que se definen como sigue.

10.1.7 Definición *Una función $f: [a, b] \to \mathbb{C}$ es **continua a trozos** si $[a, b]$ tiene una partición finita $a = x_0 < x_1 < \cdots < x_n = b$ tal que f es continua y acotada en cada intervalo abierto $]x_i, x_{i+1}[$, $i = 0, \ldots, n-1$.*

Posteriormente trabajaremos con esta clase de funciones.

10.1.8 Ejemplo *Si $f_1, \ldots, f_n$ son vectores ortonormales en el espacio con producto interno V (es decir, si $\langle f_i, f_j \rangle = 0$ para $i \neq j$ y $\langle f_i, f_i \rangle = 1$), demuéstrese que $f_1, \ldots, f_n$ son linealmente independientes.*

Solución Supóngase que $\sum_{i=1}^n c_i f_i = 0$. Debemos mostrar que $c_i = 0$. Fijando i, sea $g = \sum_{j=1}^n c_j f_j$ y formemos $\langle g, f_i \rangle$. Tenemos que

$$\langle g, f_i \rangle = \left\langle \sum_{j=1}^{n} c_j f_j, f_i \right\rangle = \sum_{j=1}^{n} c_j \langle f_j, f_i \rangle = \sum_{j=1}^{n} c_j \cdot \delta_{ji} = c_i$$

(donde $\delta_{ji} = 1$ si $j = i$ y se anula si $j \neq i$). Como $g = 0$ obtenemos $c_i = 0$, de modo que tenemos el resultado deseado. ♦

10.1.9 Ejemplo *Sea V un espacio con producto interno y $f, g \in V$, $g \neq 0$. Defínase la **proyección** de f sobre g como el vector $h = \langle f, g \rangle (g/\|g\|^2)$. Muéstrese que h y $f - h$ son ortogonales e interprétese este resultado geométricamente.*

Solución Calculamos como sigue:

$$\begin{aligned}\langle h, f - h \rangle = \langle h, f \rangle - \|h\|^2 &= \frac{\langle g, f \rangle \langle f, g \rangle}{\|g\|^2} - \frac{\langle \langle f, g \rangle g, \langle f, g \rangle g \rangle}{\|g\|^4} \\ &= \frac{\langle g, f \rangle \langle f, g \rangle}{\|g\|^2} - \frac{\langle f, g \rangle \overline{\langle f, g \rangle}}{\|g\|^2} = 0,\end{aligned}$$

pues $\langle f, g \rangle = \overline{\langle g, f \rangle}$. En consecuencia, h y $f - h$ son ortogonales. El significado geométrico de esto se muestra en la figura 10.1-2. ♦

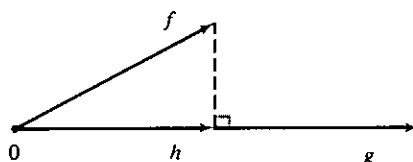


FIGURA 10.1-2 La función h es la proyección de f sobre g

Ejercicios de §10.1

1. Demuéstrese lo siguiente en un espacio con producto interno:

a. $\langle f, g\rangle = 0 \Rightarrow \|f + g\|^2 = \|f\|^2 + \|g\|^2$ (teorema de Pitágoras).

b. $4\langle f, g\rangle = (\|f+g\|^2 - \|f-g\|^2) + i(\|f+ig\|^2 - \|f-ig\|^2)$ (identidad de polarización).

c. $2\|f\|^2 + 2\|g\|^2 = \|f+g\|^2 + \|f-g\|^2$ (ley del paralelogramo).

d. $\|f+g\| \cdot \|f-g\| \leq \|f\|^2 + \|g\|^2$.

2. Muéstrese que

$$\left|\int_a^b f(x)\,dx\right|^2 \leq (b-a)\int_a^b |f(x)|^2\,dx,$$

y dedúzcase de esto que una función de cuadrado integrable en $[a, b]$, continua en $]a, b[$, también es integrable. ¿Se cumple el recíproco?

3. Sean $\varphi_1, \ldots, \varphi_n$ vectores ortonormales en V y $f \in V$. Defínase la ***proyección de*** f sobre $\varphi_1, \ldots, \varphi_n$ como $g = \sum_{i=1}^n \langle f, \varphi_i\rangle \varphi_i$. Muéstrese que g y $f - g$ son ortogonales. Interprétese geométricamente.

4. En un espacio con producto interno, demuéstrese que $\big|\,\|f\| - \|g\|\,\big| \leq \|f-g\|$. En particular, $\|f\| \leq \|g\| + \|f-g\|$. [Sugerencia: Escríbase $f = (f-g) + g$ y aplíquese la desigualdad triangular.]

5. Si $f_n \to f$ (en media), demuéstrese que $\|f_n\|$ es una sucesión acotada.

§10.2 Familias de funciones ortogonales

En esta sección estudiamos algunas propiedades generales de los vectores ortogonales en un espacio con producto interno. Los principales conceptos que desarrollaremos aquí son los de una serie de Fourier general, un sistema ortonormal completo y las relaciones entre ambos conceptos.

Sea V un espacio vectorial con producto interno $\langle\ ,\ \rangle$. Un vector $\varphi \in V$ está ***normalizado*** si $\|\varphi\| = \langle\varphi, \varphi\rangle^{1/2} = 1$. Para $g \in V$, si $g \neq 0$, entonces $g/\|g\|$ está normalizado. Además, recuérdese que f y g son ***vectores ortogonales*** si $\langle f, g\rangle = 0$.

Una sucesión $\varphi_0, \varphi_1, \varphi_2, \ldots$ en V es una ***familia ortonormal*** si cada φ_i está normalizada y φ_i, φ_j son ortogonales si $i \neq j$. Podemos abreviar estas condiciones como

$$\langle\varphi_i, \varphi_j\rangle = \delta_{ij},$$

donde $\delta_{ij} = 1$ si $i = j$ y 0 si $i \neq j$.

En última instancia, estudiaremos el espacio $V = L^2$ formado por las funciones de cuadrado integrable $f: [a, b] \to \mathbb{C}$ con $\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\,dx$, como se señaló en §**10.1**, pero por ahora nuestro estudio será en el contexto de espacios generales con producto interno.

El objeto del análisis de Fourier es escribir cada $f \in V$ en la forma

$$f = \sum_{k=0}^{\infty} c_k\varphi_k,$$

donde $c_k \in \mathbb{C}$ y $\varphi_0, \varphi_1, \varphi_2, \ldots$ es una familia ortonormal dada. En general, esto no se puede hacer; sin embargo, si se puede hacer para cada $f \in V$, la familia $\varphi_0, \varphi_1, \varphi_2, \ldots$ se denomina ***completa*** (esto no debe confundirse con el concepto de completitud que indica que todas las sucesiones de Cauchy convergen).

Se sobreentiende que la suma $\sum_{k=0}^{\infty} c_k\varphi_k$ se toma en el sentido de "convergencia en media"; es decir, si $s_n = \sum_{k=0}^{n} c_k\varphi_k$, entonces $\left\|\sum_{k=0}^{\infty} c_k\varphi_k - s_n\right\| \to 0$ cuando $n \to \infty$. Las cuestiones de convergencia puntual o uniforme (en el caso en que V sea un espacio de funciones) se tratarán en secciones posteriores.

Nuestra primera tarea será determinar las constantes c_k en la expresión para f. Esto es fácil si tenemos en cuenta la intuición geométrica. Nos referiremos concretamente al hecho de que, en $\mathbb{R}^n$, si $e_1, \ldots e_n$ es una base ortonormal, cada $x \in \mathbb{R}^n$ se escribe como

$$x = \sum_{i=1}^{n} x_i e_i,$$

donde $x_i = \langle x, e_i\rangle$. Este valor es la ***proyección*** de x a lo largo de e_i. Lo mismo es válido en el caso general.

10.2.1 Teorema *En un espacio con producto interno V, supóngase que $f = \sum_{k=0}^{\infty} c_k\varphi_k$ para una familia ortonormal $\varphi_0, \varphi_1, \ldots$ en V (convergencia en media) y que $f \in V$. Entonces $c_k = \langle f, \varphi_k\rangle = \overline{\langle\varphi_k, f\rangle}$.*

Reunimos alguna terminología importante en la siguiente definición.

10.2.2 Definición *Una familia ortonormal $\varphi_0, \varphi_1, \ldots$ en un espacio con producto interno V es **completa** si cada $f \in V$ se puede escribir como $f = \sum_{k=0}^{\infty} c_k \varphi_k$. Decimos que $\sum_{k=0}^{\infty} \langle f, \varphi_k \rangle \varphi_k$ es la **serie de Fourier** de f con respecto a $\varphi_0, \varphi_1, \ldots$ y que $\langle f, \varphi_k \rangle$ son sus **coeficientes de Fourier.***

El teorema **10.2.1** dice que el único candidato para representar f en términos de φ_k es la serie de Fourier, es decir $c_k = \langle f, \varphi_k \rangle$. Decir que $\varphi_0, \varphi_1, \ldots$ es completa es equivalente a la condición de que cada f sea "igual" a su serie de Fourier; es decir, que la serie de Fourier de f converja en media a f. Así, *$\{\varphi_k\}$ es completa sii para toda $f \in V$, $\left\| f - \sum_{k=0}^{n} \langle f, \varphi_k \rangle \varphi_k \right\| \to 0$ cuando $n \to \infty$.*

Antes de seguir adelante con la teoría, daremos algunos ejemplos de familias ortonormales (que, como veremos posteriormente, son familias *completas*).

En primer lugar, están las ***series de Fourier clásicas***, donde las φ_n son las funciones dadas por

$$\varphi_n(x) = \frac{e^{inx}}{\sqrt{2\pi}}, \quad n = 0, \pm 1, \pm 2, \ldots, \quad x \in [0, 2\pi].$$

La serie de Fourier de $f : [0, 2\pi] \to \mathbb{C}$ para esta familia está dada por

$$\sum_{k=-\infty}^{\infty} \frac{c_k e^{ikx}}{\sqrt{2\pi}}$$

donde

$$c_k = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(x) e^{-ikx}\, dx = \langle f, \varphi_k \rangle$$

(el término e^{-ikx} en la expresión de c_k tiene un signo menos, debido a que usamos el complejo conjugado de g en $\langle f, g \rangle$). Después de que demostremos la completitud (**§10.3**), podremos afirmar que f es igual a su serie de Fourier en el sentido de la convergencia en media.

Otra familia relacionada con la anterior por $e^{-inx} = \cos nx + i \operatorname{sen} nx$ y denominada también la serie de Fourier clásica está dada por

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos mx}{\sqrt{\pi}}, \frac{\operatorname{sen} nx}{\sqrt{\pi}}, \quad m, n = 1, 2, \ldots .$$

Verificaremos la ortonormalidad de esta familia en el ejemplo **10.2.8**. El lector debe escribir la serie de Fourier de una función con respecto a esta familia.

Las anteriores son las familias ortonormales más adecuadas para nuestro análisis posterior. Sin embargo, como referencia, daremos otros ejemplos clásicos que surgen en ciertas aplicaciones; para describirlas, primero daremos un repaso del ***procedimiento de Gram-Schmidt.***

Dado un espacio con producto interno V y vectores linealmente independientes g_0, $g_1, g_2, \ldots$ en V, se puede formar un sistema ortonormal correspondiente $\varphi_0, \varphi_1, \ldots$ como sigue:

$$\varphi_0 = \frac{g_0}{\|g_0\|},$$
$$\varphi_1 = \frac{g_1 - \langle g_1, \varphi_0\rangle\varphi_0}{\|g_1 - \langle g_1, \varphi_0\rangle\varphi_0\|},$$
$$\varphi_2 = \frac{g_2 - \langle g_2, \varphi_1\rangle\varphi_1 - \langle g_2, \varphi_0\rangle\varphi_0}{\|g_2 - \langle g_2, \varphi_1\rangle\varphi_1 - \langle g_2, \varphi_0\rangle\varphi_0\|},$$

etcétera. Geométricamente, esto se obtiene usando repetidamente las proyecciones ortogonales. El lector debe verificar que el proceso conduce a una familia ortonormal; véase el ejercicio 2 de esta sección.

Los ***polinomios de Legendre*** normalizados se obtienen al aplicar el procedimiento de Gram-Schmidt a los polinomios $1, x, x^2, \ldots, x^n, \ldots$, en $[-1, 1]$. Se puede demostrar por inducción (una demostración tediosa pero directa) que el n-ésimo polinomio de Legendre normalizado es

$$P_n(x) = \frac{2n+1}{\sqrt{2}\cdot 2^n n!}\frac{d^n}{dx^n}(x^2-1)^n.$$

En $\mathbb{R} =]-\infty, \infty[$, el procedimiento de Gram-Schmidt aplicado a las funciones $x^n e^{-x^2/2}$, $n = 0, 1, 2, \ldots$, da las ***funciones de Hermite*** normalizadas, mientras que aplicado a las funciones $x^n e^{-x}$, $n = 0, 1, 2, \ldots$, en $[0, \infty[$, da las ***funciones de Laguerre*** normalizadas. Estas funciones se tratan en el contexto de las ecuaciones diferenciales, donde representan soluciones fundamentales de algunas de estas ecuaciones, así como sen nx es una solución fundamental de la cuerda vibrante.

Continuaremos ahora con la teoría general.

10.2.3 Desigualdad de Bessel *Sea $\varphi_0, \varphi_1, \ldots$ un sistema ortonormal en un espacio con producto interno V. Para cada $f \in V$, la serie real $\sum_{i=0}^{\infty}|\langle f, \varphi_i\rangle|^2$ converge, y tenemos la desigualdad*

$$\sum_{i=0}^{\infty}|\langle f, \varphi_i\rangle|^2 \leq \|f\|^2.$$

Entonces, los *coeficientes de Fourier $c_k = \langle f, \varphi_k\rangle$ convergen a 0 cuando $k \to \infty$*. Así, la desigualdad de Bessel da cierto control sobre el comportamiento de los coeficientes de Fourier.

El siguiente resultado relaciona la completitud de un sistema $\varphi_0, \varphi_1, \ldots$ con la desigualdad de Bessel.

10.2.4 Teorema de Parseval *Sea V un espacio con producto interno y $\varphi_0, \varphi_1, \ldots$ un sistema ortonormal. Entonces $\varphi_0, \varphi_1, \ldots$ es completo sii para cada $f \in V$ tenemos*

$$\|f\|^2 = \sum_{n=0}^{\infty} |\langle f, \varphi_n \rangle|^2.$$

En consecuencia, vemos que la desigualdad de Bessel se convierte en una igualdad exactamente cuando se cumple la completitud. Este teorema proporciona muchas relaciones útiles en las series de Fourier, pero por lo general, no es muy práctico para decir si una familia específica $\varphi_0, \varphi_1, \ldots$ es completa. Para ello se utilizan habitualmente técnicas directas, dadas en **§10.3**.

Geométricamente, la relación de Parseval se puede considerar como un *teorema de Pitágoras generalizado*. Si g es perpendicular a h (es decir, $\langle g, h \rangle = 0$), entonces $\| g + h \|^2 = \| g \|^2 + \| h \|^2$, que es el teorema de Pitágoras para triángulos rectángulos. Si $\sum_{n=0}^{\infty} \langle f, \varphi_n \rangle \varphi_n = f$, entonces f es una suma de vectores ortogonales $\langle f, \varphi_n \rangle \varphi_n$, de modo que $\|f\|^2$ debe ser igual a la suma de los cuadrados de las longitudes de $\langle f, \varphi_n \rangle \varphi_n$. Como φ_n está normalizado, $\langle f, \varphi_n \rangle \varphi_n$ tiene como cuadrado de su longitud, $|\langle f, \varphi_n \rangle|^2$ por lo que deberíamos obtener $\|f\|^2 = \sum_{n=0}^{\infty} |\langle f, \varphi_n \rangle|^2$, que es la relación de Parseval. Si tenemos un sistema ortonormal incompleto, entonces, intuitivamente, faltarán algunos términos en el miembro derecho, por lo que sólo será válida una desigualdad, la desigualdad de Bessel.

Hemos visto que es natural y, de hecho, obligatorio, elegir los coeficientes de Fourier $c_i = \langle f, \varphi_i \rangle$ al desarrollar $f = \sum_{i=0}^{\infty} c_i \varphi_i$. Hay otra razón para esta elección que ayuda a comprenderla geométricamente. Las constantes c_i para las que la longitud

$$\left\| f - \sum_{i=0}^{n} c_i \varphi_i \right\|$$

es mínima son $c_i = \langle f, \varphi_i \rangle$, los coeficientes de Fourier (es decir, la elección $c_i = \langle f, \varphi_i \rangle$ da la *mejor aproximación media*). Esto es razonable, pues $\sum_{i=0}^{n} \langle f, \varphi_i \rangle \varphi_i$ es justamente la proyección de f en el espacio generado por $\varphi_0, \ldots, \varphi_n$, y la distancia más corta de un punto a un plano es la distancia perpendicular. Véase la figura 10.2-1. La afirmación precisa se da a continuación.

10.2.5 Teorema de la mejor aproximación media *Sea V un espacio con producto interno y $\varphi_0, \varphi_1, \ldots, \varphi_n$ un conjunto de vectores ortonormales en V. Entonces, para cada conjunto de números $t_0, t_1, \ldots, t_n$,*

$$\left\| f - \sum_{k=0}^{n} t_k \varphi_k \right\| \geq \left\| f - \sum_{k=0}^{n} \langle f, \varphi_k \rangle \varphi_k \right\|$$

La igualdad se cumple sii $t_k = \langle f, \varphi_k \rangle$.

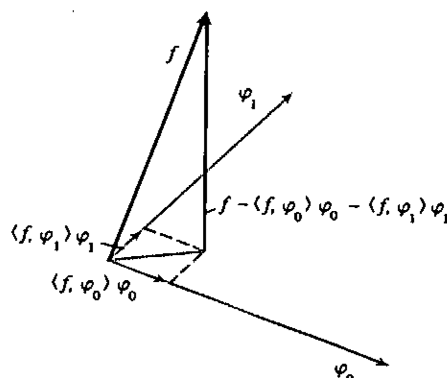


FIGURA 10.2-1 La interpretación geométrica que encierra la mejor aproximación media

Esto concluye nuestra introducción a la teoría general. Dedicamos el resto del capítulo al estudio de los casos clásicos de las familias ortonormales

$$\left\{\frac{e^{inx}}{\sqrt{2\pi}},\quad n = 0, \pm 1, \ldots\right\} \quad \text{y} \quad \left\{\frac{1}{\sqrt{2\pi}}\ \frac{\operatorname{sen} nx}{\sqrt{\pi}}, \frac{\cos mx}{\sqrt{\pi}},\ n, m = 1, 2, \ldots\right\}$$

en $[0, 2\pi]$ o $[-\pi, \pi]$. Las familias ortonormales correspondientes en otros intervalos se siguen de éstas mediante un cambio de escala (véase el ejercicio 3 de esta sección).

No podemos dejar de enfatizar que $f = \sum_{k=0}^{\infty} \langle f, \varphi_k \rangle \varphi_k$ sólo significa que la suma *converge en media* a f y que esto no implica la convergencia puntual sin ciertas condiciones adicionales. En la situación general (véase **§10.3**), obtenemos por lo general la convergencia en media, pero para la convergencia puntual o la uniforme necesitamos hipótesis más cuidadosas, como la continuidad o derivabilidad de la función f.

10.2.6 Ejemplo *Sea V un espacio con producto interno y $\varphi_0, \varphi_1, \ldots$ un sistema ortonormal completo. Muéstrese que $\varphi_1, \varphi_2, \ldots$ no es completo.*

Solución Si $\varphi_1, \varphi_2, \ldots$ fuera completo, podríamos escribir $f = \sum_{i=1}^{\infty} \langle f, \varphi_i \rangle \varphi_i$ para cada $f \in V$. Tómese $f = \varphi_0$; entonces tendríamos

$$\varphi_0 = \sum_{i=1}^{\infty} \langle \varphi_0, \varphi_i \rangle \varphi_i.$$

Pero $\langle \varphi_0, \varphi_i \rangle = 0$, $i = 1, 2, \ldots$, de modo que $\varphi_0 = 0$, lo que es imposible, pues $\|\varphi_0\| = 1$. Por lo tanto, $\varphi_1, \varphi_2, \ldots$ no es completo. Otra alternativa para resolver el problema es observar que la relación de Parseval

$$\|f\|^2 = \sum_{i=1}^{\infty} |\langle f, \varphi_i \rangle|^2$$

no es válida para $f = \varphi_0$, pues el miembro izquierdo sería 1, mientras que el miembro derecho sería 0. Análogamente, ni $\varphi_N, \varphi_{N+1}, \ldots$ ni ninguna otra subcolección *propia* es completa. ♦

10.2.7 Ejemplo *Si $\varphi_0, \varphi_1, \ldots$ es un sistema ortonormal completo en un espacio con producto interno V y f es ortogonal a cada φ_i, entonces $f = 0$.*

Solución Como el sistema es completo, podemos escribir

$$f = \sum_{i=0}^{\infty} \langle f, \varphi_i \rangle \varphi_i.$$

Por hipótesis, cada $\langle f, \varphi_i \rangle = 0$, de modo que $f = 0$ (si V fuera un espacio de Hilbert, el recíproco también sería cierto; véase el ejercicio 14 al final de este capítulo). ♦

10.2.8 Ejemplo *Muéstrese que las funciones*

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{\operatorname{sen} nx}{\sqrt{\pi}}, \quad \frac{\cos mx}{\sqrt{\pi}}, \quad n, m = 1, 2, \ldots,$$

son ortonormales en $[0, 2\pi]$.

Solución En efecto, este problema significa que

$$\int_0^{2\pi} \left(\frac{1}{\sqrt{2\pi}} \right)^2 dx = 1, \quad \int_0^{2\pi} \frac{\cos^2 mx}{\pi} dx = 1, \quad \int_0^{2\pi} \frac{\operatorname{sen}^2 nx}{\pi} dx = 1$$

(para la normalización); y (para la ortogonalidad),

$$\int_0^{2\pi} \frac{1}{\sqrt{2\pi}} \cos mx \, dx = 0,$$

$$\int_0^{2\pi} \frac{1}{\sqrt{2\pi}} \operatorname{sen} nx \, dx = 0,$$

$$\int_0^{2\pi} \frac{1}{\pi} \cos mx \operatorname{sen} nx \, dx = 0,$$

para todo $m,n \geq 1$;

$$\int_0^{2\pi} \frac{1}{\pi} \cos mx \cos m'x\, dx = 0, \;\; m \neq m';$$

y

$$\int_0^{2\pi} \frac{1}{\pi} \operatorname{sen} nx \operatorname{sen} n'x\, dx = 0, \;\; n \neq n'.$$

Cada una de estas relaciones se puede verificar mediante técnicas del cálculo para la integración de funciones trigonométricas. Una forma más fácil de hacer esto es observar que

$$\frac{1}{2\pi} \int_0^{2\pi} e^{inx} e^{-imx}\, dx = \delta_{nm},$$

puesto que si $n \neq m$,

$$\int_0^{2\pi} e^{i(n-m)x}\, dx = \left. \frac{1}{i(n-m)} e^{i(n-m)x} \right|_0^{2\pi} = 0.$$

Al obtener las partes reales e imaginarias de esta relación para todo n,m, nos dan las relaciones deseadas.

Este método muestra también que

$$\left\{ \frac{e^{inx}}{\sqrt{2\pi}} \;\middle|\; n = 0, \pm 1, \pm 2, \ldots \right\}$$

es un sistema ortonormal en $[0, 2\pi]$. ♦

10.2.9 Ejemplo *Sea $f: [0, 2\pi] \to \mathbb{C}$ tal que*

$$\int_0^{2\pi} |f(x)|^2\, dx < \infty.$$

Muéstrese que

$$\lim_{n\to\infty} \int_0^{2\pi} f(x) \operatorname{sen} nx\, dx = 0 \qquad y \qquad \lim_{m\to\infty} \int_0^{2\pi} f(x) \cos mx\, dx = 0.$$

Solución Por el ejemplo **10.2.8**, los conjuntos

$$\left\{ \frac{\operatorname{sen} nx}{\sqrt{\pi}} \;\middle|\; n = 1, 2, \ldots \right\} \quad \text{y} \quad \left\{ \frac{\cos mx}{\sqrt{\pi}} \;\middle|\; m = 1, 2, \ldots \right\}$$

son familias ortonormales. En consecuencia, por la desigualdad de Bessel, los coeficientes de Fourier de f con respecto a estos sistemas convergen a cero, de donde se sigue el resultado.

El lector puede preguntar, con justa razón, dónde se usa la hipótesis

$$\|f\|^2 = \int_0^{2\pi} |f(x)|^2 \, dx < \infty$$

en esta solución; esto se necesita para formar el espacio con producto interno $\mathcal{V}$ de tales funciones y obtener la cota superior $\|f\|^2 < \infty$, de modo que la serie en la desigualdad de Bessel converja, y entonces el n-ésimo término tenderá a cero. Más adelante regresaremos a un estudio más cuidadoso de estos puntos en el curso de nuestra demostración de los teoremas de completitud (como el lema de Riemann-Lebesgue). ♦

Ejercicios de §10.2

1. Muéstrese que cualesquiera n vectores ortonormales en $\mathbb{R}^n$ forman un conjunto completo.

2. Sean $g_0, g_1, g_2, \ldots$ vectores linealmente independientes en un espacio con producto interno. Defínase inductivamente

$$h_0 = g_0, \quad \varphi_0 = \frac{h_0}{\|h_0\|}, \ldots, \quad h_n = g_n - \sum_{k=0}^{n-1} \langle g_n, \varphi_k \rangle \varphi_k, \varphi_n = \frac{h_n}{\|h_n\|}, \ldots.$$

Muéstrese que $\varphi_0, \varphi_1, \varphi_2, \ldots$ son ortonormales. ¿Por qué debemos suponer que los g_i son linealmente independientes?

3. **a.** Supóngase que $\varphi_0(x), \varphi_1(x), \ldots$ son funciones ortonormales en $[0, 2\pi]$. Muéstrese que las funciones

$$\psi_n(x) = \sqrt{\frac{2\pi}{l}}\, \varphi_n\left(\frac{2\pi x}{l}\right)$$

son ortonormales en $[0, l]$.

b. Escríbase la familia obtenida al adaptar $1/\sqrt{2\pi}$, $(\operatorname{sen} nx)/\sqrt{\pi}$, $(\cos mx)/\sqrt{\pi}$ o, alternativamente, $e^{inx}/\sqrt{2\pi}$, a $[0, l]$ como en **a.**

c. Escríbase la serie de Fourier de f para las familias obtenidas en **b.**

d. Muéstrese que si las funciones φ_n en **a** son completas, también lo son las ψ_n.

4. Supóngase por el momento que las funciones $1/\sqrt{2\pi}$, $(\operatorname{sen} nx)/\sqrt{\pi}$, $(\cos mx)/\sqrt{\pi}$ son completas en el intervalo $[0, 2\pi]$ (demostraremos esto posteriormente).

a. Aplíquese esto a la función x para mostrar que $x = \pi - 2\sum_{n=1}^{\infty}(\operatorname{sen} nx)/n$ (convergencia en media).

b. Úsense los coeficientes de Fourier determinados en **a** y aplíquese la relación de Parseval para mostrar que $\pi^2/6 = \sum_{n=1}^{\infty} 1/n^2$.

c. Úsese el mismo procedimiento con x^2 para obtener $\pi^4/90 = \sum_{n=1}^{\infty} 1/n^4$.

5. Demuéstrese que

$$2\sum_{k=1}^{n} \cos k\theta = \frac{\operatorname{sen}[(n+1/2)\theta]}{\operatorname{sen}(\theta/2)} - 1$$

usando $e^{i\theta} + e^{2i\theta} + \cdots + e^{ni\theta} = \dfrac{e^{i\theta}(1 - e^{in\theta})}{1 - e^{i\theta}}$.

6. Demuéstrese que la serie de Fourier de una suma de dos funciones es la suma término a término de sus series de Fourier.

§10.3 Completitud y teoremas de convergencia

Esta sección estudia la convergencia de la serie de Fourier de una función. La serie de Fourier de una función queda determinada de forma única por dicha función, pero no existe una garantía *a priori* de que la serie converja, o bien, si converge, que su suma sea la función dada. El tipo de convergencia que obtenemos depende de las hipótesis acerca de f. Los resultados importantes se resumen en la tabla 10.3-1; en **§10.4** y **§10.6** daremos más teoremas de convergencia.

Es posible debilitar un poco las hipótesis de los teoremas presentados sobre la convergencia puntual; algunas de esas hipótesis ligeramente más precisas aparecen en la sección opcional **§10.4.**[1] Los ejemplos y técnicas de cálculo aparecen en **§10.5.**

[1] Existe un resultado profundo de L. Carleson que postula que para $|f|^2$ integrable, la serie de Fourier de f converge puntualmente a f, excepto posiblemente en un conjunto de medida nula. Sin embargo, este teorema está más allá del objetivo de este libro. Véase el *Acta Math.* **116** (1966), pág. 135.

TABLA 10.3-1 Propiedades de convergencia de las series de Fourier

Hipótesis sobre la función f	Convergencia de la serie de Fourier
$\int_0^{2\pi} \lvert f(x)\rvert^2\, dx < \infty$	Converge en media a f
f, f' ambas continuas a trozos, y con sólo discontinuidades de salto	Converge puntualmente (y en media) a $\dfrac{f(x^+)+f(x^-)}{2}$
f continua, con $f(-\pi) = f(\pi)$, f' continua a trozos y con sólo discontuidades de salto	Converge uniformemente (§**10.6**), puntualmente y en media a f

A partir de este momento, trabajaremos principalmente con los siguientes dos sistemas ortonormales en el intervalo $[0, 2\pi]$ o $[-\pi, \pi]$:

a. Sistema exponencial:

$$\frac{e^{inx}}{\sqrt{2\pi}}, \qquad n = 0, \pm 1, \pm 2, \ldots;$$

b. Sistema trigonométrico:

$$\frac{1}{\sqrt{2\pi}}, \frac{\operatorname{sen} nx}{\sqrt{\pi}}, \frac{\cos mx}{\sqrt{\pi}}, \qquad n, m = 1, 2, \ldots.$$

Estos dos sistemas están íntimamente relacionados; de hecho, el sistema trigonométrico se obtiene como las partes real e imaginaria del sistema exponencial (véase el ejercicio 1 de esta sección).

La serie de Fourier de una función $f: [0, 2\pi] \to \mathbb{C}$ con respecto del sistema exponencial es la serie

$$\sum_{n=-\infty}^{\infty} c_n e^{inx} = \lim_{N\to\infty} \sum_{n=-N}^{N} c_n e^{inx},$$

donde los coeficientes de Fourier están dados por

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} f(y) e^{-iny}\, dy$$

(hemos reunido dos $\sqrt{2\pi}$ para obtener 2π por conveniencia, así como por razones históricas y convencionales).

La serie de Fourier de una función f con respecto del sistema trigonométrico es

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \operatorname{sen} nx,$$

donde los coeficientes están dados por

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos nx \, dx, \quad n = 0, 1, 2, \ldots$$

y

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \operatorname{sen} nx \, dx, \quad n = 1, 2, \ldots.$$

El lector debe repasar **§10.1** y **§10.2** si estos postulados no han quedado claros.

Las sumas parciales de la serie trigonométrica y de la serie exponencial son iguales (ejercicio 1 al final de esta sección). Así, si podemos demostrar los teoremas de convergencia para una de estas series, automáticamente obtendremos los teoremas para la otra. El sistema utilizado depende del problema particular y, en cierta medida, los gustos personales. En **§10.5** se dan algunos ejemplos de las diferencias en los cálculos.

Nuestro objetivo principal es dar teoremas que indiquen cuándo una función es "igual" a su serie de Fourier. Si entendemos la igualdad en el sentido de la convergencia en media, entonces éste es un problema de completitud del sistema ortonormal. Por fortuna, y éste es uno de los principales teoremas del tema, los sistemas exponencial y trigonométrico son completos. Por otro lado, si entendemos la igualdad en el sentido de la convergencia puntual o uniforme, se necesitan más condiciones sobre f. Primero trataremos la completitud:

10.3.1 Teorema de completitud media *Los sistemas exponencial y trigonométrico en* $[0, 2\pi]$ *(o* $[-\pi, \pi]$*) son completos en el espacio* $V = L^2$ *de funciones* $f: [0, 2\pi] \to \mathbb{C}$ *con* $\int_0^{2\pi} |f(x)|^2 dx < \infty$ *(la integral puede ser impropia).*

Esto significa que si una función f tiene $|f|^2$ integrable (es decir, la función f es cuadrado integrable), entonces f es igual a la suma de su serie de Fourier, en el sentido de convergencia en media. No necesariamente se obtiene la convergencia puntual. De hecho, uno puede construir una función continua (periódica) f cuya serie de Fourier diverja en un punto dado (véase, por ejemplo, Widom, Drasin y Tromba, *Lectures on Measure and Integration Theory*, Van Nostrand Mathematical Studies, núm. 20, Nueva York, 1967, pág. 153).

Sea s_n la n-ésima suma parcial de la serie de Fourier trigonométrica de una función f con valores reales. La idea intuitiva detrás del teorema de completitud media se muestra en la figura 10.3-1. Cada s_n es una función suave (pues es un polinomio trigonométrico), pero cuando $n \to \infty$ s_n puede converger a algo discontinuo, como vimos en el capítulo 5. Así, si f es discontinua, obtenemos convergencia en media, pero *no* convergencia uniforme.

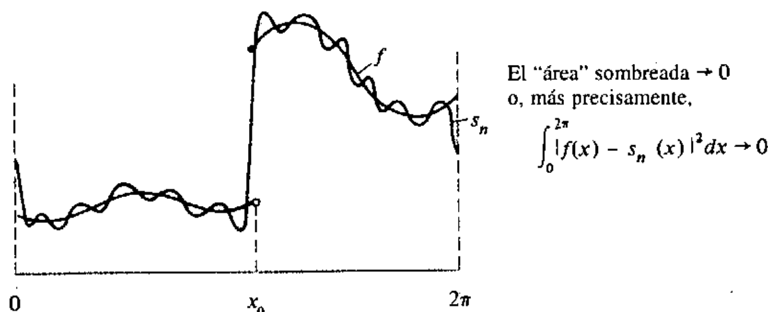


FIGURA 10.3-1—El significado geométrico de la convergencia en media

Aunque el teorema de completitud media trata una amplia variedad de funciones, no responde a la pregunta sobre la convergencia puntual. El siguiente teorema analiza esta cuestión. Para establecer el teorema, necesitamos cierta terminología. Para este teorema podemos usar funciones reales o complejas, pero basta considerar solamente las funciones reales, ya que si $f = f_1 + if_2$, la serie de Fourier de f es la de $f_1 + i$ (la de f_2) (¿por qué?). Supóngase que $f : [0, 2\pi] \to \mathbb{R}$ (o $[-\pi, \pi] \to \mathbb{R}$) tiene una posible discontinuidad en $x_0 \in [0, 2\pi]$. Si $x_0 = 0$ o 2π, esto significaría que estamos tomando la función f *extendida periódicamente*; es decir, definimos $f(x + 2\pi) = f(x)$. Esto es razonable, pues la propia serie de Fourier es periódica. Esta extensión periódica se ilustra para dos casos en la figura 10.3-2. Definimos

$$f(x_0^+) = \lim_{x \to x_0^+} f(x) = \lim_{x \to x_0,\ x > x_0} f(x)$$

si existe. Esto significa que para cada $\varepsilon > 0$, existe $\delta > 0$ tal que si $|x - x_0| < \delta$ y $x > x_0$, entonces $|f(x) - f(x_0^+)| < \varepsilon$. Intuitivamente, $f(x_0^+)$ indica el valor de f justo a la derecha de x_0. Véase la figura 10.3-3. Por supuesto, $f(x_0^+)$ podría no existir; obsérvese la figura 10.3-3b. Se define $f(x_0^-)$ de forma completamente análoga. Una discontinuidad en x_0 tal que $f(x_0^-)$ y $f(x_0^+)$ existen se denomina ***discontinuidad de salto*** y $f(x_0^+) - f(x_0^-)$ recibe el nombre de ***salto*** de f en x_0. El salto puede ser positivo o negativo, y se anula cuando $f(x_0) = f(x_0^+) = f(x_0^-)$ sii f es continua en x_0.

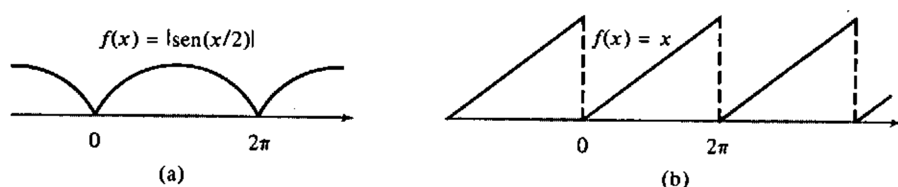


FIGURA 10.3-2 (a) Continua en 0; (b) discontinua en 0

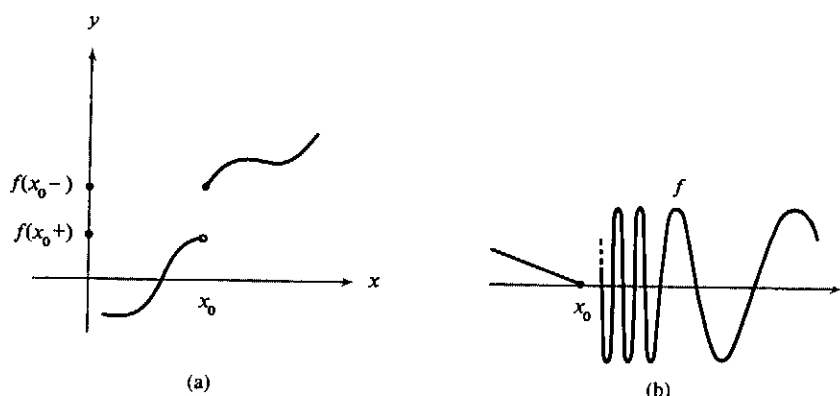


FIGURA 10.3-3 (a) Discontinuidad de salto; (b) $f(x_0^+)$ no existe

Supóngase que f es derivable en algún intervalo abierto $]x_0, x_0 + \varepsilon[$. Entonces podemos hablar de $f'(x_0^+)$ y $f'(x_0^-)$ si existen. Intuitivamente, $f'(x_0^+)$ es la pendiente de *f justo a la derecha de* x_0. Por ejemplo, en la figura 10.3-2a, $f'(0^+) = \lim_{x\to 0^+}(d/dx)(\operatorname{sen}(x/2)) = 1/2$ y en la figura 10.3-2b, $f'(0^+) = +1$.

Existe una definición un poco más débil de $f'(x_0^+)$ que a veces es importante. La definición anterior requería que $f'(x)$ existiera para $x > x_0$ y que $f'(x_0^+) = \lim_{x\to x_0^+} f'(x)$ existiera. Es fácil demostrar que si esto es así, entonces

$$f'(x_0^+) = \lim_{h\to 0^+} \left\{ \frac{f(x_0 + h) - f(x_0^+)}{h} \right\}$$

(véase el ejercicio 39 al final del capítulo). Para el siguiente teorema basta la existencia del límite, por lo que la adoptaremos como nuestra definición, donde $f'(x_0^-)$ se define de forma análoga como

$$f'(x_0^-) = \lim_{h\to 0^+} \left\{ \frac{f(x_0^-) - f(x_0 - h)}{h} \right\}.$$

La demostración de que este segundo método es en realidad una hipótesis más débil se deja al lector en el mismo ejercicio. Obsérvese que f es derivable en x_0 sii $f(x_0) = f(x_0^+) = f(x_0^-)$ y $f'(x_0^+)$ y $f'(x_0^-)$ existen y son iguales.

El siguiente teorema (debido a Jordan) contiene el resultado principal relativo a la convergencia puntual.

10.3.2 Teorema de la convergencia puntual *Sea $f: [0, 2\pi] \to \mathbb{R}$ (o $f: [-\pi, \pi] \to \mathbb{R}$) continua a trozos, con una discontinuidad de salto en x_0, y supóngase que $f'(x_0^+)$ y $f'(x_0^-)$ existen. Entonces la serie de Fourier de f (en forma exponencial o trigonométrica) evaluada en x_0 converge a $[f(x_0^+) + f(x_0^-)]/2$. En particular, si f es derivable en x_0, la serie de Fourier de f converge en x_0 a $f(x_0)$.*

Si x_0 es un extremo del intervalo, entonces, como mencionamos anteriormente, los números $f(x_0^+)$ y $f(x_0^-)$ de la función se calculan después de extenderla en forma periódica (véase la figura 10.3-2). Obsérvese también que la serie de Fourier no necesariamente converge a $f(x_0)$ en una discontinuidad de salto, sino que converge al *promedio* de $f(x_0^+)$ y $f(x_0^-)$. Un ejemplo típico es una función escalonada (véase la figura 10.3-4).

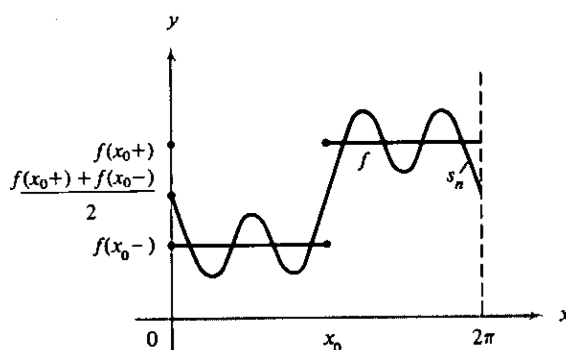


FIGURA 10.3-4 En una discontinuidad, la serie de Fourier converge al valor promedio

El teorema de la convergencia puntual es útil, pues nos da condiciones que se verifican fácilmente en los ejemplos y que son válidas en muchos casos de interés. Sin los beneficios del teorema de la convergencia puntual, sería difícil, incluso en los ejemplos sencillos, demostrar directamente que la serie de Fourier converge a la función asociada.

En el teorema de la convergencia puntual también tenemos la convergencia en media de la serie de Fourier a f, por el teorema de completitud media. Sin embargo, la convergencia puntual es una condición más delicada y, con frecuencia, más útil.

10.3.3 Ejemplo *Supóngase que* $f: [0, 2\pi] \to \mathbb{C}$ *satisface* $\int_0^{2\pi} |f(x)|^2 dx < \infty$. *Muéstrese que*

$$\int_0^{2\pi} |f(x)|^2\, dx = \sum_{n=-\infty}^{\infty} \frac{1}{2\pi} \left| \int_0^{2\pi} f(x) e^{-inx}\, dx \right|^2$$

$$= \frac{1}{2\pi} \left| \int_0^{2\pi} f(x)\, dx \right|^2 + \sum_{n=1}^{\infty} \frac{1}{\pi} \left| \int_0^{2\pi} f(x) \cos nx\, dx \right|^2$$

$$+ \sum_{n=1}^{\infty} \frac{1}{\pi} \left| \int_0^{\pi} f(x) \operatorname{sen} nx\, dx \right|^2 .$$

Solución Sea

$$\mathcal{V} = L^2 = \left\{ f : [0, 2\pi] \to \mathbb{C} \;\middle|\; \|f\|^2 = \int_0^{2\pi} |f(x)|^2\, dx < \infty \right\}.$$

Por el teorema de completitud media, tenemos la familia ortonormal completa

$$\left\{ \varphi_n(x) = \frac{e^{inx}}{\sqrt{2\pi}} \;\middle|\; n = 0, \pm 1, \pm 2, \ldots \right\}.$$

Así, es válida la relación de Parseval: $\|f\|^2 = \sum_{n=-\infty}^{\infty} |\langle f, \varphi \rangle|^2$, donde

$$\langle f, \varphi_n \rangle = \int_0^{2\pi} f(x) \overline{\varphi_n(x)}\, dx = \int_0^{2\pi} \frac{f(x) e^{-inx}\, dx}{\sqrt{2\pi}},$$

de modo que se sigue la primera igualdad. Recuérdese que $\sum_{n=-\infty}^{\infty} |\langle f, \varphi_n \rangle|^2$ significa que $\lim_{N \to \infty} \sum_{n=-N}^{N} |\langle f, \varphi_n \rangle|^2$ para las funciones exponenciales (es decir, se toman en el orden $\varphi_0, \varphi_{\pm 1}, \varphi_{\pm 2}, \ldots$; en este caso, los términos son positivos, de modo que la serie se puede reordenar de forma arbitraria, por el ejemplo resuelto **5.5** al final del capítulo 5).

La segunda igualdad es consecuencia de la aplicación del mismo procedimiento a la familia ortonormal completa

$$\left\{ \frac{1}{\sqrt{2\pi}}, \frac{\operatorname{sen} nx}{\sqrt{\pi}}, \frac{\cos mx}{\sqrt{\pi}}, \quad n, m = 1, 2, \ldots \right\}.$$

Esta segunda igualdad también se puede obtener de la primera escribiendo $e^{-inx} = \cos nx - i \operatorname{sen} nx$, elevando al cuadrado y reagrupando términos y observando que los términos cruzados de n y $-n$ se cancelan. ♦

10.3.4 Ejemplo *Para las siguientes funciones en* $[-\pi, \pi]$, *establezca si la serie de Fourier converge en media o puntualmente y a lo que converge en* $x_0 = 0$.

a. $f(x) = \begin{cases} -2 & x \leq 0 \\ 2 & x > 0. \end{cases}$

b. $f(x) = \begin{cases} x & x \leq 1 \\ x+1 & x > 1. \end{cases}$

c. $f(x) = \operatorname{sen} x.$

d. $f(x) = \begin{cases} 1+x & x \leq 0 \\ x + \operatorname{sen}(1/x) & x > 0. \end{cases}$

Solución En la figura 10.3-5 aparecen las gráficas de estas funciones. Cada función es continua a trozos y las discontinuidades son discontinuidades de salto. Esto es evidente, excepto posiblemente para **d**. En ese caso, $f(x) = x \operatorname{sen}(1/x) \to 0$ cuando $x \to 0$, pues $|x \operatorname{sen}(1/x)| \leq |x|$, de modo que $f(0^+) = 0$.

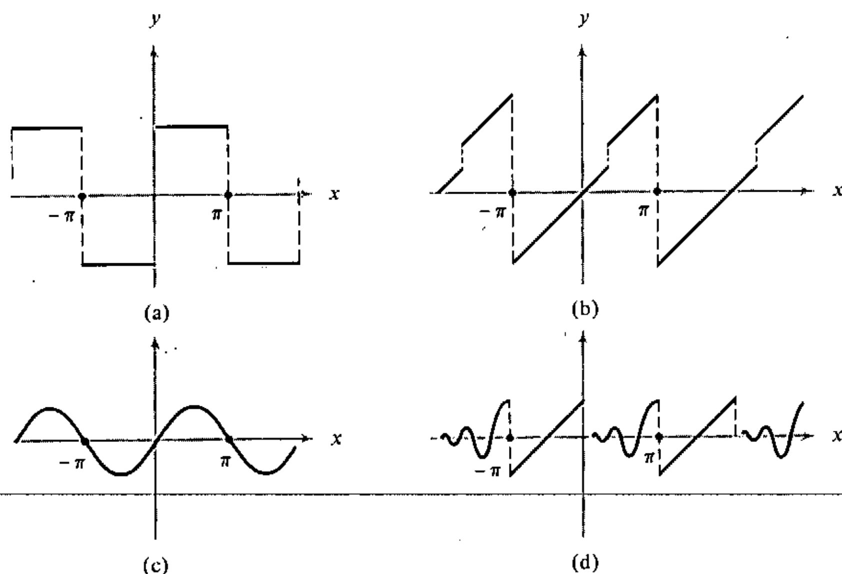


FIGURA 10.3-5 Determínense las propiedades de convergencia de la serie de Fourier de estas funciones

En 0, $f'(0^+)$ y $f'(0^-)$ existen en los casos **a, b** y **c**. Todo esto es bastante obvio. Por ejemplo, en **a**, $f(x) = 2$ para $x > 0$, de modo que $\lim_{x\to 0^+} f'(x)$ existe y es igual a 0. En el caso **d** esto no es cierto, pues si $h > 0$, tenemos

$$\frac{f(0+h) - f(0^+)}{h} = \operatorname{sen}\frac{1}{h},$$

que no converge cuando $h \to 0$. Así, el teorema de convergencia puntual no se aplica a este caso. Sin embargo, en cada caso tenemos convergencia en media, por el teorema de completitud media. En $x = 0$, la serie de Fourier converge en **a** a $0 = [f(0^+) + f(0^-)]/2$, en **b** a 0, en **c** a 0 y en **d** falla el teorema (uno puede mostrar la convergencia de la serie de Fourier en **d** a 1/2 mediante un análisis directo). ♦

10.3.5 Ejemplo *Encuéntrese un ejemplo de una función f tal que la serie de Fourier de f converja puntualmente y en media a f, pero no uniformemente.*

Solución Sea

$$f = \begin{cases} 0 & -\pi < x < 0 \\ 1/2 & \text{si } x = 0 \text{ o } x = \pi \\ 1 & 0 < x < \pi. \end{cases}$$

Las discontinuidades de f son discontinuidades de salto (véase la figura 10.3-6). Por el teorema de completitud media, la serie de Fourier de f converge a f en media, y por el teorema de convergencia puntual, converge puntualmente, pues $f(x_0) = [f(x_0^+) + f(x_0^-)]/2$ en cada punto.

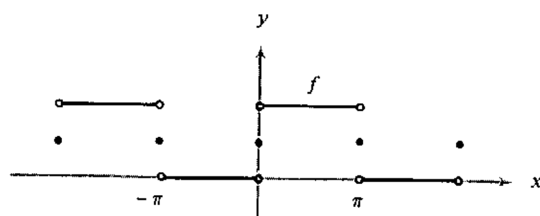


FIGURA 10.3-6 Esta función tiene una serie de Fourier que converge en media y puntualmente pero no uniformemente

Sin embargo, la serie de Fourier no puede converger uniformemente a f, pues cada $s_n(x)$ es continua, y si $s_n(x) \to f(x)$ uniformemente, f sería continua, lo cual no ocurre. ♦

Ejercicios de §10.3

1. **a.** Muéstrese que la n-ésima suma parcial de la serie de Fourier trigonométrica de una función (real o compleja) es igual a la n-ésima suma parcial de la serie exponencial.

b. Escríbase la serie correspondiente en $[-\pi, \pi]$.

c. Muéstrese que si f es par en $[-\pi, \pi]$ (es decir, $f(x) = f(-x)$), entonces todos los $b_n = 0$ en la serie de Fourier trigonométrica. La serie se llama entonces ***serie de cosenos.***

d. Repítase la pregunta **c** para f impar; es decir, si $f(-x) = -f(x)$, muéstrese que todos los $a_n = 0$. La serie se llama entonces ***serie de senos***.

2. Para $f: [0, 2\pi] \to \mathbb{R}$, muéstrese que f es continua en cero (entendida como una función periódica) sii $f(0) = f(2\pi)$ y f es continua en el sentido usual en ambos puntos 0 y 2π en $[0, 2\pi]$ (es decir, $\lim_{x\to 0^+} f(x) = f(0)$ y $\lim_{x\to 2\pi^-} f(x) = f(2\pi)$.

3. Supóngase que $f: [0, l] \to \mathbb{C}$ satisface $\int_0^l |f(x)|^2 dx < \infty$. Muéstrese que

$$\int_0^l |f(x)|^2\, dx = \frac{1}{l}\sum_{n=-\infty}^{\infty}\left|\int_0^l f(x)e^{2\pi inx/l}\, dx\right|^2$$
$$= \frac{1}{l}\left|\int_0^l f(x)\, dx\right|^2 + 2\sum_{n=1}^{\infty}\frac{1}{l}\left|\int_0^l f(x)\cos\frac{2\pi nx}{l}\, dx\right|^2$$
$$+ 2\sum_{m=1}^{\infty}\frac{1}{l}\left|\int_0^l f(x)\operatorname{sen}\frac{2\pi mx}{l}\, dx\right|^2$$

4. Úsense los teoremas de esta sección para justificar los cálculos en el ejercicio 4 de **§10.2.**

5. Para cada una de las siguientes funciones en $[-\pi, \pi]$ determínese si la serie de Fourier converge puntualmente o en media y cuál es el límite puntual, si éste existe.

a. $f(x) = x^n$ (considérense todos los valores posibles de n: . . . , –3, –2, –1, 0, 1, 2, . . .).

b. $f(x) = \begin{cases} 0 & x < 0 \\ kx & x \geq 0 \end{cases}$ para algún $k \in \mathbb{R}$.

c. $f(x) = \tan x$.

d. $f(x) = e^{-x^2}$.

e. $f(x) = \begin{cases} e^{-1/x^2} & x > 0 \\ 0 & x \leq 0. \end{cases}$

§10.4 Funciones de variación acotada y teoría de Fejér (opcional)

Existe un teorema similar al de la convergencia puntual, pero que es válido en condiciones más generales y que también da un criterio para la convergencia uniforme. Postularemos este teorema sin demostración (la demostración es similar a la del teorema de convergencia puntual, sólo que un poco más intrincada). Nos conformaremos con demostrar una versión más débil en **§10.6** y un teorema relacionado con él, debido a Fejér.

Para comprender el teorema, necesitamos el concepto de una función de variación acotada. Si $f: [a, b] \to \mathbb{R}$, decimos que f es de ***variación acotada*** si existe un número M tal que para toda partición $a = x_0 < x_1 < \cdots < x_n = b$ de $[a, b]$,

$$\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \leq M.$$

Intuitivamente, decir que f es de variación acotada significa que la *gráfica de f tiene una longitud de arco finita*. Se puede mostrar que una función es de variación acotada sii es la diferencia de dos funciones monótonas acotadas.[2] Se sigue que si f es de variación acotada, entonces todas sus discontinuidades son de salto y forman un conjunto numerable, de modo que $f(x_0^+)$ y $f(x_0^-)$ están definidos.

10.4.1 Teorema de Dirichlet-Jordan

Sea $f: [0, 2\pi] \to \mathbb{R}$ una función acotada.

i. *Si f es de variación acotada en un intervalo $[x_0 - \varepsilon, x_0 + \varepsilon]$ (para algún $\varepsilon > 0$) en torno a x_0, entonces la serie de Fourier de f evaluada en x_0 converge a $[f(x_0^+) + f(x_0^-)]/2$.*

ii. *Si f es continua y de variación acotada, entonces la serie de Fourier de f converge uniformemente a f.*

---

[2] Si f es de variación acotada, sea $v(x) = \sup\{\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \; |a = x_0 \leq x_1 \leq \cdots \leq x_n = x\}$, la variación de f. Escribimos $f = p - q$, donde $p = v + f/2$ y $q = v - f/2$. Se puede comprobar que p y q son crecientes. El recíproco es fácil de verificar.

Tanto el teorema de convergencia puntual como el teorema de Dirichlet-Jordan dan condiciones suficientes para que converja la serie de Fourier. El ejercicio 34 al final del capítulo da un ejemplo para mostrar que las condiciones no son necesarias. No se conocen condiciones necesarias y suficientes.

Como ya hemos observado, la serie de Fourier de una función continua no necesariamente converge puntualmente. Por el teorema de Dirichlet-Jordan, una tal función no puede ser de variación acotada. La teoría de Fejér analiza este caso, debilitando la necesidad de la convergencia puntual de la serie a la sumabilidad Cesaro de la serie. Recuérdese de **§5.10** que una sucesión $a_1, a_2 \ldots$ converge ***en el sentido de Cesaro***, o $(C, 1)$, si $\sigma_n = (a_1 + \cdots + a_n)/n$ converge. Si $a_n \to x$ entonces $\sigma_n \to x$, pero no necesariamente se cumple el recíproco. Para las series, este criterio se aplica a las sumas parciales. En 1904, Fejér demostró el *importante hecho* de que, aunque la serie de Fourier de una función continua no necesariamente converge puntualmente, siempre es convergente $(C,1)$.

10.4.2 Teorema de Fejér *Supóngase que f es continua a trozos en* $[0, 2\pi]$ *y que* $f(x_0^+)$ *y* $f(x_0^-)$ *existen. Entonces la serie de Fourier de f converge* $(C, 1)$ *en* x_0 *a* $[f(x_0^+) + f(x_0^-)]/2$. *Si f es continua y* $f(0) = f(2\pi)$, *la serie de Fourier converge* $(C, 1)$ *uniformemente a f.*

Obsérvese que no se requiere ninguna hipótesis de variación acotada ni de derivabilidad.

Al considerar las "distribuciones" o las "funciones generalizadas", como la función delta de Dirac (**§8.9**), la serie de Fourier sigue teniendo sentido con una interpretación adecuada, y toda distribución tiene una serie de Fourier convergente (convergente en un sentido adecuado). Estos asuntos relativos a la convergencia son muy útiles en la práctica, pero por cuestiones de espacio no nos es posible realizar aquí un tratamiento sistemático de ellos. Sin embargo, podemos dar un ejemplo.

10.4.3 Ejemplo *Calcúlese la serie de Fourier de la función delta* δ *en* $[-\pi, \pi]$.

Solución Recuérdese que esta función tiene la siguiente propiedad que la define:

$$\int_{-\pi}^{\pi} f(x)\delta(x-a)\,dx = f(a).$$

En consecuencia:

$$\int_{-\pi}^{\pi} \delta(x)e^{-inx}\,dx = 1,$$

de modo que la serie de Fourier de δ es

$$\sum_{n=-\infty}^{\infty} \frac{e^{inx}}{2\pi}.$$

Por supuesto, esta serie no converge en $x = 0$, pero tampoco lo esperábamos, pues $\delta(0)$ no está definida. Lo que sí es cierto es que

$$\delta(x) = \sum_{n=-\infty}^{\infty} \frac{e^{inx}}{2\pi}$$

en el sentido de que la *convergencia es válida bajo el signo de integral*; es decir, para cualquier función C^1, f,

$$f(0) = \int_{-\pi}^{\pi} \delta(x)f(x)\,dx = \sum_{n=-\infty}^{\infty} \int_{-\pi}^{\pi} \frac{e^{inx}}{2\pi} f(x)\,dx$$

La validez de esto es consecuencia del teorema de la convergencia puntual:

$$f(y) = \sum_{n=-\infty}^{\infty} \left(\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx\right) \frac{e^{iny}}{2\pi}$$

para cada y (como la suma es de $-\infty$ a $+\infty$, podemos reemplazar n por $-n$). La situación para una distribución general y la demostración de la convergencia de su serie de Fourier es análoga. Es decir, si T es una distribución, entonces

$$T = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} a_n e^{inx},$$

donde $a_n = T(e^{-inx})$. ♦

Ejercicios de §10.4

1. Calcúlese la serie de Fourier de δ', la derivada de la función delta.

2. Demuéstrese que $\sum_{k=-\infty}^{\infty} e^{ikx}$ es sumable $(C, 1)$ a 0 si x no es múltiplo de 2π.

§10.5 Cálculo de series de Fourier

En esta sección nos ocuparemos de ejemplos específicos de series de Fourier y los métodos para calcularlas. En este análisis incluimos un fenómeno interesante que aparece en el comportamiento de una serie de Fourier en una discontinuidad de salto, conocido como fenómeno de Gibbs.

Las formas trigonométrica y exponencial de una serie de Fourier son equivalentes, como ya hemos visto (ejercicio 1, **§10.3**). Las diversas formas de las series de Fourier y sus propiedades de convergencia se resumen en las tablas 10.5-1 y 10.5-2. Las funciones pueden ser reales o complejas, pero trabajaremos con funciones reales para simplificar nuestra exposición. Debemos hacer varios comentarios acerca de estas fórmulas. Las dos primeras formas (tabla 10.5-1) deberían ser autoexplicativas. La ***serie de senos de Fourier*** surge cuando f es impar, pues entonces tenemos $f(-x) = -f(x)$ y, en consecuencia,

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx = \frac{1}{\pi}\int_{-\pi}^{\pi} f(-x)\cos(-nx)\,d(-x) \\ &= -\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,d(-x) = -a_n, \end{aligned}$$

de modo que $a_n = 0$.

Análogamente, si f es par, la serie de Fourier se reduce a la serie de cosenos. Véase la figura 10.5-1.

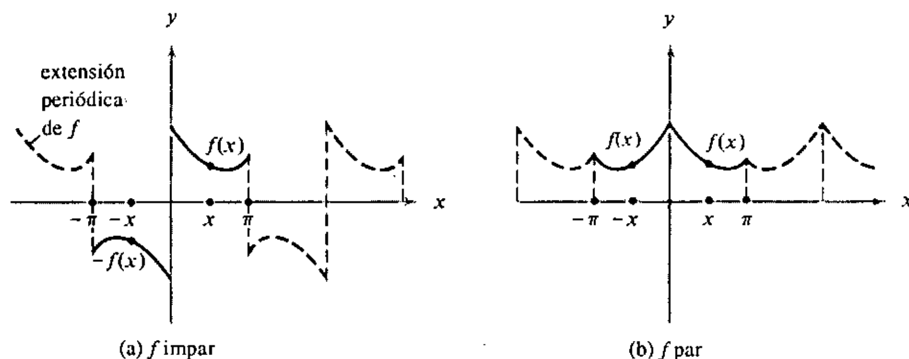


FIGURA 10.5-1 (a) f es impar; (b) f es par

Para el intervalo $[-l, l]$, reemplazamos las funciones ortonormales φ_k en $[-\pi, \pi]$ por las funciones $\psi_k(x) = \sqrt{\pi/l}\ \varphi_k(\pi x/l)$, que son ortonormales en $[-l, l]$. Este cambio de

TABLA 10.5-1 Varias formas de las series de Fourier

	La función f	Serie de Fourier	Coeficientes
Serie de Fourier exponencial	f definida en $[0, 2\pi]$ (o $[-\pi, \pi]$)	$\sum_{n=-\infty}^{\infty} c_n e^{inx}$	$c_n = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}\,dx$ (o $\int_{-\pi}^{\pi}$)
Serie de Fourier trigonométrica	f definida en $[-\pi, \pi]$ (o $[0, 2\pi]$)	$\frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos\, nx + b_n \operatorname{sen} nx)$	$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos\, nx\,dx \quad n = 0, 1, \ldots$ $b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\operatorname{sen} nx\,dx \quad n = 1, 2, \ldots$ (o $\int_0^{2\pi}$)
Serie de Fourier de senos	f definida en $[-\pi, \pi]$ y f impar; $f(-x) = -f(x)$	$\sum_{n=1}^{\infty} b_n \operatorname{sen} nx$	$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\operatorname{sen} nx\,dx \quad n = 1, 2, \ldots$ $= \frac{2}{\pi}\int_0^{\pi} f(x)\operatorname{sen} nx\,dx$
Serie de Fourier de cosenos	f definida en $[-\pi, \pi]$ y f par; $f(-x) = f(x)$	$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\, nx$	$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos\, nx\,dx \quad n = 1, 2, \ldots$ $= \frac{2}{\pi}\int_0^{\pi} f(x)\cos\, nx\,dx$
Serie exponencial en $[-l, l]$	f definida en $[-l, l]$	$\sum_{n=-\infty}^{\infty} a_n e^{inx\pi/l}$	$c_n = \frac{1}{2l}\int_{-l}^{l} f(x)e^{-inx\pi/l}\,dx$
Serie trigonométrica en $[-l, l]$	f definida en $[-l, l]$	$\frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos\, \frac{n\pi x}{l} + b_n \operatorname{sen} \frac{n\pi x}{l}\right)$	$a_n = \frac{1}{l}\int_{-l}^{l} f(x)\cos\, \frac{n\pi x}{l}\,dx \quad n = 0, 1, \ldots$ $b_n = \frac{1}{l}\int_{-l}^{l} f(x)\operatorname{sen} \frac{n\pi x}{l}\,dx \quad n = 1, 2, \ldots$
Serie de cosenos en semiintervalos	f definida en $[0, l]$	$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\, \frac{n\pi x}{l}$	$a_n = \frac{2}{l}\int_0^{l} f(x)\cos\, \frac{n\pi x}{l}\,dx \quad n = 0, 1, \ldots$
Serie de senos en semiintervalos	f definida en $[0, l]$	$\sum_{n=1}^{\infty} b_n \operatorname{sen} \frac{n\pi x}{l}$	$b_n = \frac{2}{l}\int_0^{l} f(x)\operatorname{sen} \frac{n\pi x}{l}\,dx \quad n = 1, 2, \ldots$

TABLA 10.5-2 Propiedades de convergencia

Propiedades de f	Convergencia de la serie de Fourier
$\int_0^{2\pi} \lvert f(x)\rvert^2\,dx < \infty$	Converge en media a f
f tiene una discontinuidad de salto en x_0 y $f'(x_0^+)$, $f'(x_0^-)$ existen. Si x_0 es un extremo, consideramos a f como extendida, de modo que sea periódica (véase el texto para las formas en los semiintervalos).	Converge puntualmente en x_0 a $\dfrac{f(x_0^+)+f(x_0^-)}{2}$
f continua, con $f(-\pi) = f(\pi)$, f' continua a trozos con discontinuidades de salto únicamente	Converge uniformemente a f

escala preserva las propiedades de convergencia. El lector debe escribir las series de senos y de cosenos (para f impar o par) en un intervalo arbitrario $[-l, l]$.

Las fórmulas para semiintervalos se obtienen como sigue. Para la serie de cosenos, extendemos f a $[-l, l]$ como

$$f(-x) = f(x).$$

Entonces f se convierte en una función par, por lo que tiene una serie de cosenos. Véase la figura 10.5-2. Para la convergencia en $x_0 = 0$, debemos verificar esta función extendida, no la original. Si $f(0^+)$ existe, entonces está claro que la función par extendida no tiene un salto en 0. Análogamente, no tenemos un salto en l o en $-l$. Así, aplicamos el criterio usual de convergencia sin modificar la serie de cosenos.

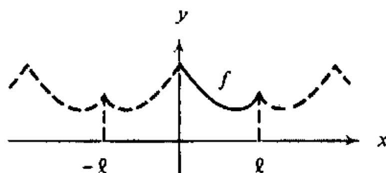


FIGURA 10.5-2 Extensión de f para que sea par

La serie de senos para semiintervalos es análoga. En $[-l, 0[$, definimos f como $f(-x) = -f(x)$, para $0 < x \leq l$, de modo que f sea impar y, por lo tanto, tenga una serie de senos

como serie de Fourier. Véase la figura 10.5-3. En este caso, la serie de Fourier siempre se anula en el punto 0 (sen 0 = 0). Así, generalmente aparece una discontinuidad de salto en 0, pero la serie de Fourier se anula en estos puntos. Para garantizar la continuidad de la f extendida, tendríamos que imponer las condiciones de que f sea continua y que $f(0) = f(l) = 0$. Así, aplicamos el criterio de convergencia a la serie de senos en el semiintervalo sin modificación alguna, si recordamos que la convergencia es a 0 en 0 y en l.

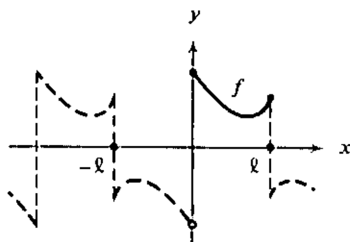


FIGURA 10.5-3 Extensión de f para que sea impar

Por la teoría general, sabemos que la relación de Parseval es válida para cada f con

$$\int_0^{2\pi} |f|^2\, dx < \infty;$$

es decir,

$$\int_0^{2\pi} |f|^2\, dx = \sum_{n=-\infty}^{\infty} |c_n|^2,$$

donde c_n son los coeficientes de Fourier en forma exponencial. Hay que tener cuidado, ya que c_n, a_n, b_n de la tabla **10.5-1** no son los coeficientes de Fourier en el sentido anterior, pues hemos reagrupado los factores $\sqrt{2\pi}, \sqrt{\pi}$ por tradición. Pero si recordamos esto, es fácil encontrar la relación de Parseval. Los resultados aparecen en la tabla 10.5-3.

Si conocemos la serie de Fourier de $f(x)$, digamos, en $[-\pi, \pi]$, entonces podemos obtener un desarrollo de la función $g(x) = \int_{-\pi}^{x} f(y)dy$ por medio del siguiente teorema.

10.5.1 Teorema de integración *Supóngase que* $\int_{-\pi}^{\pi} |f(x)|^2 dx < \infty$ *y que* f *tiene la serie de Fourier*

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \operatorname{sen} nx).$$

TABLA 10.5-3 Relación de Parseval

Tipo de serie	Relación de Parseval
Serie exponencial	$\frac{1}{2\pi}\int_0^{2\pi} \lvert f(x)\rvert^2\,dx = \sum_{n=-\infty}^{\infty} \lvert c_n\rvert^2 \left(\text{o } \int_{-\pi}^{\pi}\right)$
Serie trigonométrica	$\frac{1}{\pi}\int_{-\pi}^{\pi} \lvert f(x)\rvert^2\,dx = \frac{a_0^2}{2} + \sum_{n=1}^{\infty}(a_n^2 + b_n^2) \left(\text{o } \int_0^{2\pi}\right)$
Serie de senos	$\frac{1}{\pi}\int_{-\pi}^{\pi} \lvert f(x)\rvert^2\,dx = \sum_{n=1}^{\infty} b_n^2$
Serie de cosenos	$\frac{1}{\pi}\int_{-\pi}^{\pi} \lvert f(x)\rvert^2\,dx = \frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2$
Serie exponencial en $[-l, l]$	$\frac{1}{l}\int_{-l}^{l} \lvert f(x)\rvert^2\,dx = \sum_{n=-\infty}^{\infty} \lvert c_n\rvert^2$
Serie trigonométrica en $[-l, l]$	$\frac{1}{l}\int_{-l}^{l} \lvert f(x)\rvert^2\,dx = \frac{a_0^2}{2} + \sum_{n=1}^{\infty}(a_n^2 + b_n^2)$
Serie de cosenos en semiintervalos	$\frac{2}{l}\int_0^{l} \lvert f(x)\rvert^2\,dx = \frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2$
Serie de senos en semiintervalos	$\frac{2}{l}\int_0^{l} \lvert f(x)\rvert^2\,dx = \sum_{n=1}^{\infty} b_n^2$

Entonces, si $g(x) = \int_{-\pi}^{x} f(y)dy$, *tenemos*

$$g(x) = \frac{a_0(x+\pi)}{2} + \sum_{n=1}^{\infty}\left(a_n\int_{-\pi}^{x}\cos ny\,dy + b_n\int_{-\pi}^{x}\operatorname{sen} ny\,dy\right)$$

$$= \frac{a_0(x+\pi)}{2} + \sum_{n=1}^{\infty}\left\{\frac{a_n}{n}\operatorname{sen} nx + \frac{b_n}{n}\left((-1)^n - \cos\, nx\right)\right\}$$

y la convergencia es uniforme para $-\pi \le x \le \pi$.

Obsérvese que este desarrollo no es la serie de Fourier de g, sino que da el desarrollo de Fourier de $g(x) - a_0x/2$. Además, la expresión se obtiene simplemente al integrar la serie de Fourier de f término a término. Análogamente, cualquiera de las series de la tabla 10.5-1 se puede integrar término a término para obtener una serie uniformemente convergente (la demostración es la misma en cada caso). Esto es útil cuando las constantes a_n y b_n correspondientes a f ya se han calculado. Para obtener la serie de Fourier de g, sustituimos la serie de x y agrupamos términos (la serie de x aparece posteriormente en esta sección; véase también el ejemplo resuelto **10.3** al final del capítulo).

Como en §**5.3**, la derivación de series de Fourier requiere más cuidado que la integración; la analizaremos en §**10.6.**

En la tabla 10.5-4 reunimos algunos de los desarrollos de Fourier más comunes. Al usar esta tabla, debemos recordar que la serie de Fourier es lineal; es decir,

$$(\text{serie de Fourier de } af + bg) = a(\text{serie de Fourier de } f) + b(\text{serie de Fourier de } g).$$

También hay que recordar que si se encuentra un desarrollo en funciones trigonométricas, éste debe ser la serie de Fourier, pues ésta es única.

El teorema **10.5.1** se puede usar para construir más series, integrando sucesivamente, por ejemplo, $x, x^2, x^3, \ldots$. Además, obsérvese que si f se modifica en un número finito de puntos (o incluso una cantidad numerable), la serie de Fourier no se modifica (¿por qué?). Veremos algunos ejemplos específicos más adelante. En la tabla 10.5-4, se presentan intervalos de longitud 2π y π por conveniencia, los cuales se pueden cambiar por intervalos de longitud $2l$ y l, introduciendo constantes y nuevas variables, como se indica en la tabla 10.5-1. En los ejercicios y ejemplos aparecen más desarrollos.

En estas fórmulas, hay que tener cuidado con el dominio. Por ejemplo, $f(x) = x$ en $[0, 2\pi]$ es un poco diferente de $f(x) = x$ en $[-\pi, \pi]$ como función periódica (figura 10.5-4).

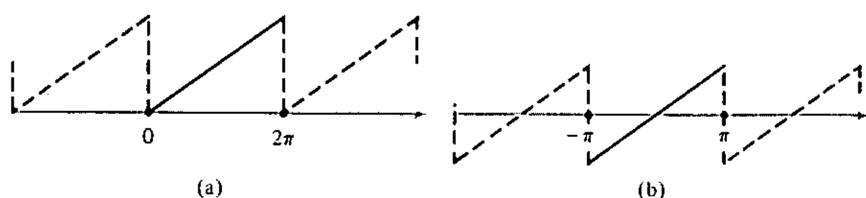


FIGURA 10.5-4 (a) $f(x) = x$ en $[0, 2\pi]$; (b) $f(x) = x$ en $[-\pi, \pi]$

Por supuesto, en $]0, \pi[$, las funciones son iguales puntualmente a sus series. La comparación de estas series produce identidades interesantes. Por ejemplo, para la función x, deducimos que

$$x = \pi - 2\sum_{n=1}^{\infty} \frac{\operatorname{sen} nx}{n} = 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \operatorname{sen} nx,$$

TABLA 10.5-4 Algunas series de Fourier y otras series relacionadas

Función	Serie	Válida puntualmente en el intervalo
1 $f(x) = \begin{cases} 1 & 0 \le x \le \pi \\ 0 & -\pi \le x < 0 \end{cases}$	$\frac{1}{2} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\operatorname{sen}[(2n-1)x]}{2n-1} = \frac{1}{2} + \frac{1}{\pi} \sum_{n=1}^{\infty} (1-(-1)^n) \frac{\operatorname{sen} nx}{n}$	$]-\pi, \pi[$; la serie tiene el valor $\frac{1}{2}$ en $x = 0, \pi, -\pi$
1a $f(x) = 1,\ 0 \le x \le \pi$	$\frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\operatorname{sen}[(2n-1)x]}{2n-1}$ (serie de senos en un semiintervalo)	$]0, \pi[$; la serie tiene el valor 0 en $x = 0, x = \pi$
	1 (serie de cosenos en un semiintervalo)	$[0, \pi]$
2 $f(x) = x$	$2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \operatorname{sen} nx$	$]-\pi, \pi[$; la serie tiene el valor 0 en $x = \pi, x = -\pi$
	$\pi - 2 \sum_{n=1}^{\infty} \frac{\operatorname{sen} nx}{n}$	$]0, 2\pi[$; la serie tiene el valor π en $x = 0, x = 2\pi$
	$\frac{\pi}{2} - \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\cos[(2n-1)x]}{(2n-1)^2}$ (serie de cosenos en un semiintervalo)	$[0, \pi]$

Tabla 10.5-4 *Continuación.*

2a $f(x) = \begin{cases} 0 & -\pi \le x < 0 \\ x & 0 \le x \le \pi \end{cases}$	$\frac{\pi}{4} - \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\cos[(2n-1)x]}{(2n-1)^2} - \sum_{n=1}^{\infty}\frac{(-1)^n}{n}\operatorname{sen} nx$	$]-\pi, \pi[$; la serie tiene el valor $\pi/2$ en $x = \pi, x = -\pi$
3 $f(x) = x^2$	$2\pi x - \frac{2\pi^2}{3} + 4\sum_{n=1}^{\infty}\frac{\cos nx}{n^2}$	$[0, 2\pi]$
	$\frac{4}{3}\pi^2 + 4\sum_{n=1}^{\infty}\left(\frac{\cos nx}{n^2} - \pi\frac{\operatorname{sen} nx}{n}\right)$	$]0, 2\pi[$; la serie tiene el valor $2\pi^2$ en $0, 2\pi$
	$\frac{\pi^2}{3} + 4\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos nx$	$[-\pi, \pi]$
	$\pi x - \frac{8}{\pi}\sum_{n=1}^{\infty}\frac{\operatorname{sen}[(2n-1)x]}{(2n-1)^3}$	$[0, \pi]$
	$\sum_{n=1}^{\infty}\left(\frac{2\pi(-1)^{n+1}}{n} - \frac{4}{\pi}\frac{1-(-1)^n}{n^3}\right)\operatorname{sen} nx$ (serie de senos en un semiintervalo)	$[0, \pi[$; la serie tiene el valor 0 en $x = \pi$
4 $f(x) = \operatorname{sen} x$	$\frac{2}{\pi} - \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\cos 2nx}{4n^2-1}$ (serie de cosenos en un semiintervalo)	$[0, \pi[$

Tabla 10.5-4 *Continuación.*

4a $f(x) = \begin{cases} \operatorname{sen} x & 0 \le x \le \pi \\ 0 & -\pi \le x < 0 \end{cases}$	$\frac{1}{\pi} + \frac{1}{2} \operatorname{sen} x - \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\cos 2nx}{4n^2 - 1}$	$[-\pi, \pi]$
4b $f(x) = \lvert \operatorname{sen} x \rvert$	$\frac{2}{\pi} - \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\cos 2nx}{4n^2 - 1}$	todo $\mathbb{R}$
5 $f(x) = \cos x$	$\frac{8}{\pi} \sum_{n=1}^{\infty} \frac{n \operatorname{sen} 2nx}{4n^2 - 1}$ (serie de senos en un semiintervalo)	$]0, \pi[$; la serie tiene el valor 0 en $x = 0, x = \pi$
6 $f(x) = e^x$	$\frac{\operatorname{senh} \pi}{\pi} \sum_{-\infty}^{\infty} \frac{(-1)^n}{1 - in} e^{inx}$	$]-\pi, \pi[$; la serie tiene el valor $\cosh \pi$ en $\pi, -\pi$

por lo que recuperamos **1a** en la tabla 10.5-4:

$$\sum_{n=1}^{\infty} \frac{\operatorname{sen}[(2n-1)x]}{2n-1} = \frac{\pi}{4}$$

para $0 < x < \pi$. Sin embargo, fuera del intervalo $]0, \pi[$ *no* coinciden; véase la figura 10.5-5. A continuación damos un bosquejo de la apariencia de ambas series hasta el n-ésimo término.

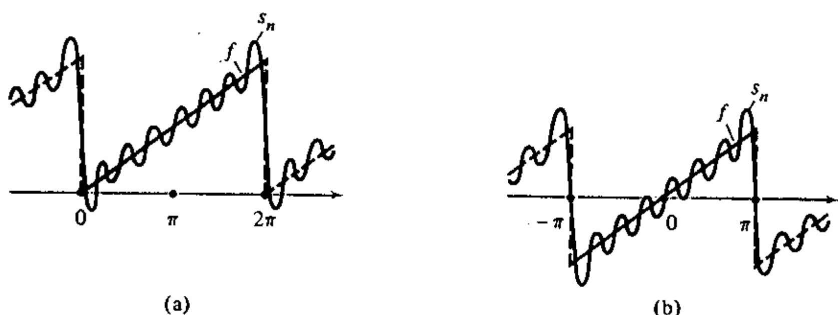


FIGURA 10.5-5 La serie de Fourier de (a) $f(x) = x$ en $[0, 2\pi]$ y (b) en $[-\pi, \pi]$

Analizaremos ahora el ***fenómeno de Gibbs***, que ocurre por lo general cuando f tiene una discontinuidad de salto (recibe el nombre de J. Gibbs, físico-matemático y físico-químico que lo descubrió; usualmente, Gibbs también recibe reconocimiento por el desarrollo de la notación vectorial hacia 1880). La idea se ilustra en la figura 10.5-6. Esto muestra que si s_n es la n-ésima suma parcial de la serie de Fourier, entonces el máximo y el mínimo de s_n cerca del salto están más separados que el salto de f, y este exceso se mantiene cuando $n \to \infty$. Intuitivamente, la serie de Fourier "sobreestima" el salto, y esta sobreestimación continúa hasta en el límite. Otra forma de decir esto es que cuando $n \to \infty$, $s_n(x)$ tiende a aproximarse a una recta vertical más larga que el salto.

Nos conformaremos con tratar un caso particular de la discontinuidad de salto para el cual podemos calcular explícitamente la sobreestimación.

10.5.2 Fenómeno de Gibbs *Considérese*

$$f(x) = \begin{cases} a & -\pi \le x < 0 \\ b & 0 \le x \le \pi, \end{cases}$$

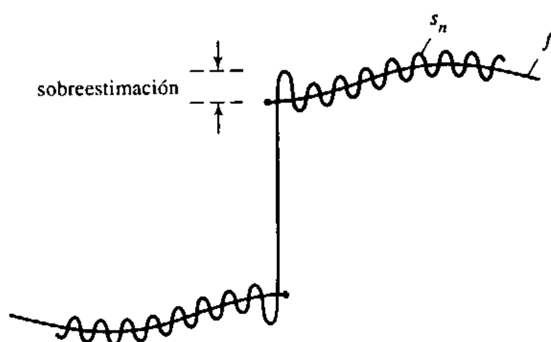


FIGURA 10.5-6 El fenómeno de Gibbs

y supóngase que $a < b$. Sea $s_n(x)$ la n-ésima suma parcial de la serie de Fourier trigonométrica. Entonces el máximo de s_n se encuentra en $\pi/2n$ y el mínimo en $-\pi/2n$, y, además,

$$\lim_{n\to\infty} s_n\left(\frac{\pi}{2n}\right) = \left(\frac{b-a}{2}\right)\left(\frac{2}{\pi}\int_0^{\pi}\frac{\operatorname{sen} t}{t}\,dt + 1\right) + b \approx (b-a)(0.089) + b.$$

Análogamente,

$$\lim_{n\to\infty} s_n\left(-\frac{\pi}{2n}\right) = \left(\frac{b-a}{2}\right)\left(-\frac{2}{\pi}\int_0^{\pi}\frac{\operatorname{sen} t}{t}\,dt + 1\right) + a \approx a - (b-a)(0.089),$$

y la diferencia de estos límites es

$$\left(\frac{b-a}{2}\right)\left(-\frac{4}{\pi}\int_0^{\pi}\frac{\operatorname{sen} t}{t}\,dt + 1\right) \approx (b-a)(1.179).$$

Para $a = -1$ y $b = 1$, lo cual se ilustra en la figura 10.5-7. Así, la sobreestimación del máximo y el mínimo es aproximadamente del 9 por ciento del salto de f en cada caso.

10.5.3 Ejemplo *Verifíquese la fórmula* **1** *de la tabla 10.5-4.*

Solución La serie se obtiene al evaluar los coeficientes de Fourier mediante la integración directa en $[-\pi, \pi]$ (véase la tabla 10.5-1). Obtenemos

$$a_n = \frac{1}{\pi}\int_0^{\pi}\cos nx\,dx = \begin{cases} 1 & n = 0 \\ 0 & n = 1, 2, \ldots \end{cases}$$

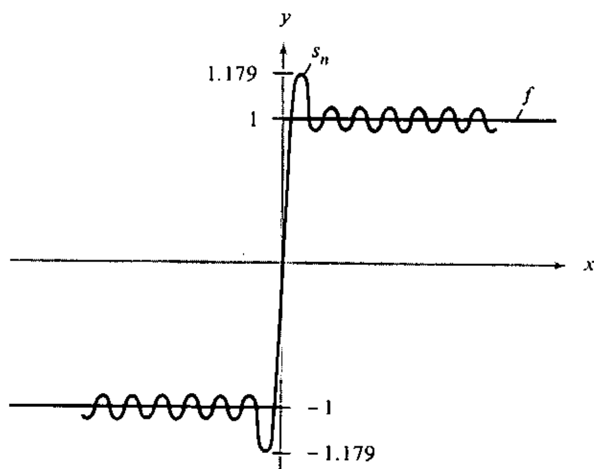


FIGURA 10.5-7 El fenómeno de Gibbs para una función escalón

y

$$b_n = \frac{1}{\pi}\int_0^{\pi} \operatorname{sen} nx\, dx = \frac{1}{\pi}\left(-\frac{\cos nx}{n}\right)\Bigg|_0^{\pi} = \frac{1}{\pi}\left(\frac{1-(-1)^n}{n}\right),$$

que es igual a $2/\pi n$ si n es impar y a cero si n es par. Esto establece la fórmula 1. ♦

10.5.4 Ejemplo *Úsense las tablas para determinar la serie correspondiente a la función*

$$g(x) = \begin{cases} -1 & -\pi \le x < 0 \\ 1 & 0 \le x \le \pi . \end{cases}$$

Solución Sea f dada en la fórmula 1 de la tabla 10.5-4. Entonces

$$g(x) = 2f(x) - 1,$$

de modo que el desarrollo de g es

$$2\left(\frac{1}{2} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\operatorname{sen}[(2n-1)x]}{2n-1}\right) - 1 = \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\operatorname{sen}[(2n-1)x]}{2n-1},$$

pues el desarrollo de Fourier de 1 en $[-\pi, \pi]$ es 1 (esto también se puede obtener en forma directa). ♦

El *desarrollo de senos* de 1 en un semiintervalo no es el propio 1, sino el desarrollo de g en el ejemplo **10.5.4** (¿por qué?).

10.5.5 Ejemplo *Para cada* $0 < x < \pi$, *demuéstrese que*

$$\frac{\pi}{2} - \sum_{n=1}^{\infty} \frac{2}{\pi n^2}(1 - (-1)^n)\cos nx = 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \operatorname{sen} nx.$$

Solución El miembro izquierdo es la serie de cosenos para $f(x) = x$ en $[0, \pi]$, mientras que el miembro derecho es la serie de senos. Para cada $0 < x < \pi$ tenemos convergencia al valor x, de modo que obtenemos la identidad deseada (en $\pm\pi$, el miembro derecho se anula; ¿cuánto vale el miembro izquierdo?). ♦

10.5.6 Ejemplo *Establézcase la primera fórmula del apartado 2 de la tabla 10.5-4 para* $f(x) = x$ *y determínese la forma de obtener las fórmulas para* x^2.

Solución Como $f(x) = x$ es impar, usamos la serie de senos. Entonces

$$b_n = \frac{2}{\pi}\int_0^{\pi} x \operatorname{sen} nx \, dx.$$

Una integración por partes da

$$b_n = \frac{2}{\pi}\left(-\frac{x\cos nx}{n}\right)\Big|_0^{\pi} + \frac{2}{\pi}\int_0^{\pi} \frac{\cos nx}{n}\,dx$$
$$= \frac{2}{\pi}\frac{\pi(-1)^{n+1}}{n} + \frac{2}{\pi}\cdot 0.$$

Por lo tanto, la serie es

$$\sum_{n=1}^{\infty} b_n \operatorname{sen} nx = 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \operatorname{sen} nx.$$

La serie de x^2 se obtiene como sigue. La primera fórmula es la integral de la serie de x en $[0, 2\pi]$ (con un factor 2, pues $\int_0^x y\,dy = x^2/2$). El término $2\pi^2/3$ proviene del término del coseno en 0, usando $\sum_{n=1}^{\infty} 1/n^2 = \pi^2/6$ (ejercicio 4, §10.2). Aquí hemos usado el teorema **10.5.1**. La segunda fórmula utiliza la primera y el desarrollo de x en $]0, 2\pi[$ de **2** en la tabla 10.5-4 (las primeras dos fórmulas también se pueden encontrar directamente). La tercera fórmula es la serie de Fourier (= serie de cosenos, en este caso) para x^2. Las demás fórmulas se obtienen calculando la integral del desarrollo en cosenos de x en $[0, \pi]$ y usando el desarrollo en senos de x en $]0, \pi[$. ♦

10.5.7 Ejemplo *Determínese la serie de Fourier en $[-\pi, \pi]$ para*

$$f(x) = \begin{cases} 0 & -\pi \le x \le 0 \\ x^2 & 0 \le x \le \pi. \end{cases}$$

Solución Si integramos la función **2a** de la tabla 10.5-4, obtenemos 1/2 de la función f dada aquí. En consecuencia, por el teorema **10.5.1**,

$$\frac{1}{2} f(x) = \frac{\pi x}{4} + \frac{\pi^2}{4} - \frac{2}{\pi} \sum_{n=1}^{\infty} \int_{-\pi}^{x} \frac{\cos[(2n-1)x]}{(2n-1)^2} dx$$
$$- \sum_{n=1}^{\infty} \int_{-\pi}^{x} \frac{(-1)^n}{n} \operatorname{sen} nx\, dx$$
$$= \frac{\pi x}{4} + \frac{\pi^2}{4} - \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\operatorname{sen}[(2n-1)x]}{(2n-1)^3}$$
$$+ \sum_{n=1}^{\infty} (-1)^n \left(\frac{\cos nx}{n^2}\right) - \frac{(-1)^n}{n^2} \cdot (-1)^n.$$

Si insertamos la primera serie de $f(x) = x$ de **2** obtenemos

$$\frac{\pi}{4}\left(2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \operatorname{sen} nx\right) + \frac{\pi^2}{4} - \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\operatorname{sen}[(2n-1)x]}{(2n-1)^3}$$
$$+ \sum_{n=1}^{\infty} \frac{(-1)^n \cos nx}{n^2} - \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

Ésta es la serie deseada, aunque aún hay que pulirla. Como (ejercicio 4, §**10.2**) $\sum_{n=1}^{\infty} 1/n^2 = \pi^2/6$, la serie resultante es

$$f(x) = 2\left(\frac{\pi^2}{12} + \sum_{n=1}^{\infty} \left\{ \left[\frac{\pi(-1)^{n+1}}{2n} - \frac{1-(-1)^n}{\pi n^3}\right] \operatorname{sen} nx + \frac{(-1)^n}{n^2} \cos nx \right\}\right). \quad \blacklozenge$$

Ejercicios de §10.5

1. Establézcase lo siguiente en la tabla 10.5-4:
   a. Las fórmulas **2, 2a**
   b. Las fórmulas **4, 4a, 4b**

2. Establézcase lo siguiente:

$$x\cos x = -\frac{1}{2}\operatorname{sen} x + 2\sum_{n=2}^{\infty}\frac{(-1)^n n \operatorname{sen} nx}{n^2-1}, \quad -\pi < x < \pi.$$

3. Determínense las series de senos y de cosenos en un semiintervalo para x^2.

4. Calcúlese la serie de Fourier en $[-\pi, \pi]$ para cada una de las siguientes funciones:

a. $f(x) = \begin{cases} 10 & x > 0 \\ -11 & x < 0 \end{cases}$

b. $f(x) = \begin{cases} x & x > 0 \\ -x^2 & x < 0 \end{cases}$

c. $f(x) = x^2 + x + 3$

5. Calcúlese el salto en el fenómeno de Gibbs para la función

$$f(x) = \begin{cases} -8 & x > 0 \\ -4 & x < 0 \end{cases}$$

en el intervalo $[-\pi, \pi]$.

6. Considerando

$$f(x) = \begin{cases} \pi/4 & 0 \le x \le \pi \\ -\pi/4 & -\pi \le x < 0 \end{cases}$$

y el punto $x = \pi/2$, demuéstrese la *fórmula de Leibniz*:

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \frac{\pi}{4}.$$

§10.6 Más teoremas de convergencia

En esta sección daremos algunos teoremas adicionales de convergencia relativos principalmente a la convergencia uniforme, la derivabilidad y la integración de las series de Fourier. En **§10.4** establecimos que si f es continua y de variación acotada, entonces la serie de Fourier de f converge uniformemente a f. Daremos ahora una versión un poco más débil que es más sencilla de demostrar, y que resulta casi igual de útil en la práctica. El lector debería repasar el concepto de convergencia uniforme del capítulo 5. Por ejemplo, en la figura 10.5-6, ¿por qué s_n no es uniformemente convergente a f?

10.6.1 Teorema de convergencia uniforme *Supóngase que f es continua en $[-\pi, \pi]$, $f(-\pi) = f(\pi)$ y que f' es continua a trozos, con discontinuidades de salto. Entonces la serie de Fourier (trigonométrica o exponencial) de f converge a f absoluta y uniformemente. Una afirmación similar se cumple para f en* $[0, 2\pi]$.

En particular, esto implica que la serie de Fourier converge en media y puntualmente, lo que es consistente con lo aprendido en **§10.3.**

Por ejemplo, considérese la función $f(x) = |x|$ en $[-\pi, \pi]$. En este caso, las condiciones del teorema **10.6.1** se satisfacen (pero no se satisfacen en $[0, 2\pi]$, pues no es la misma función), de modo que la serie de Fourier de f, a saber,

$$\frac{\pi}{2} - \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\cos[(2n-1)x]}{(2n-1)^2},$$

converge uniformemente. Véase la figura 10.6-1.

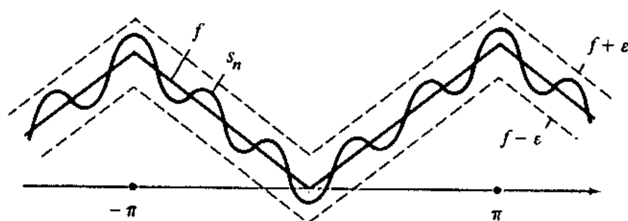


FIGURA 10.6-1 La serie de Fourier de esta función converge uniformemente

Así, por la definición de convergencia uniforme, para cada $\varepsilon > 0$ existe N tal que $n \geq N$ implica

$$\left|\, |x| - \left(\frac{\pi}{2} - \frac{4}{\pi}\sum_{1}^{n}\frac{\cos[(2n-1)x]}{(2n-1)^2}\right)\right| < \varepsilon$$

para todo $x \in [-\pi, \pi]$.

Uno podría pensar que si

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \operatorname{sen} nx),$$

entonces

$$f'(x) = \sum_{n=1}^{\infty}\{(-na_n)\operatorname{sen} nx + nb_n \cos nx\}$$

en los puntos en que $f'(x)$ existe. Por desgracia, esto no es cierto. Por ejemplo, sea

$$f(x) = \begin{cases} 1 & 0 < x \leq \pi \\ -1 & -\pi \leq x \leq 0; \end{cases}$$

entonces

$$f(x) = \frac{4}{\pi} \sum_{n=1}^{\infty} \operatorname{sen} \frac{[(2n-1)x]}{(2n-1)},$$

de modo que para $x > 0$ esperaríamos $0 = (4/\pi) \sum_{n=1}^{\infty} \cos[(2n-1)x]$. Pero esta serie no converge, pues $\cos[(2n-1)x]$ no converge a 0 (para que todo esto tenga sentido podemos usar la teoría de distribuciones).

Para obtener un teorema de derivación, naturalmente pensaríamos en usar las técnicas del capítulo 5. Sin embargo, en este caso podemos obtener un teorema mejor con un argumento directo. El resultado es el siguiente.

10.6.2 Teorema de derivación *Sea f continua en* $[-\pi, \pi]$, $f(-\pi) = f(\pi)$ *y* f' *continua a trozos, con discontinuidades de salto. Supóngase que* f'' *existe en* $x \in [-\pi, \pi]$. *Entonces la serie de Fourier para*

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \operatorname{sen} nx)$$

se puede derivar término a término en x para dar

$$f'(x) = \sum_{n=1}^{\infty} (-na_n \operatorname{sen} nx + nb_n \cos nx).$$

Además, ésta es la serie de Fourier de f'.

Como en el capítulo 5, hay que tener cuidado al derivar las series; deben cumplirse ciertas condiciones para justificar las operaciones. El resultado debe compararse con el teorema de integración **10.5.1**.

10.6.3 Ejemplo Considérese $f(x) = |x|$, donde $x \in [-\pi, \pi[$, la cual satisface las condiciones del teorema de derivación en cada $x \neq 0$. Su serie de Fourier es

$$\frac{\pi}{2} - \frac{4}{\pi} \sum_{1}^{\infty} \frac{\cos[(2n-1)x]}{(2n-1)^2}.$$

En consecuencia,

$$f'(x) = \begin{cases} 1 & 0 < x \leq \pi \\ -1 & -\pi \leq x \leq 0 \end{cases}$$

tiene la serie de Fourier

$$\frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\operatorname{sen}[(2n-1)x]}{(2n-1)},$$

lo que coincide con lo que sabemos. ♦

10.6.4 Ejemplo *Proporciónese la versión del teorema de convergencia uniforme válida en* $[-l, l]$.

Solución Queremos mostrar que si f es continua en $[-l, l]$, $f(-l) = f(l)$ y $f'(x)$ es continua a trozos con discontinuidades de salto, entonces la serie de Fourier

$$\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos\frac{n\pi x}{l}+b_n\operatorname{sen}\frac{n\pi x}{l}\right)$$

converge uniforme y absolutamente a f, donde

$$a_n=\frac{1}{l}\int_{-l}^{l}f(x)\cos\frac{n\pi x}{l}\,dx \qquad \text{y} \qquad b_n=\frac{1}{l}\int_{-l}^{l}f(x)\operatorname{sen}\frac{n\pi x}{l}\,dx$$

(véase la tabla 10.5-1). Podemos realizar la demostración mediante el *método* de demostración del teorema de convergencia uniforme, pero también podemos deducir el resultado *directamente* de este teorema como sigue. Sea $g(x) : [-\pi, \pi] \to \mathbb{R}$ dada por $g(x) = f(lx/\pi)$. Entonces

$$a_n=\frac{1}{l}\int_{-l}^{l}f(x)\cos\frac{n\pi x}{l}\,dx=\frac{1}{\pi}\int_{-\pi}^{\pi}f\left(\frac{ly}{\pi}\right)\cos ny\,dy$$

(usando $x = ly/\pi$). Así, a_n es también el coeficiente de Fourier de g, y análogamente para b_n. Como g satisface las condiciones del teorema de convergencia uniforme, tenemos que

$$\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos ny+b_n\operatorname{sen} ny)$$

converge uniforme y absolutamente a g en $[-\pi, \pi]$. Si reemplazamos y por $\pi x/l$, vemos que lo mismo es cierto para f. ♦

10.6.5 Ejemplo *Para cada una de las siguientes funciones, explíquese si la serie de Fourier converge en media, puntualmente o uniformemente. Determínese si podemos derivar la serie de Fourier.*

a. $f: [0, 2\pi] \to \mathbb{R}$, *donde* $f(x) = 1/n$ *si* $1/(n+1) \leq x < 1/n$ *y* $f(x) = 1$ *si* $1/2 \leq x \leq 2\pi$, *donde* $n = 1, 2, \ldots$.

b. $f: [-\pi, \pi] \to \mathbb{R}$, *donde* $f(x) = \pi - |x|$.

c. $f: [-\pi, \pi] \to \mathbb{R}$, *donde* $f(x) = x^2 + 1$ *si* $-\pi \leq x < 0$ *y* $f(x) = x + 1$ *si* $0 < x \leq \pi$.

Solución La gráfica de las tres funciones aparece en la figura 10.6-2. En los tres casos, f está acotada y es de cuadrado integrable (la función en **a** es integrable, pues sus discontinuidades forman un conjunto numerable; véase el capítulo 8). En consecuencia, la serie de Fourier converge en media en todos los casos.

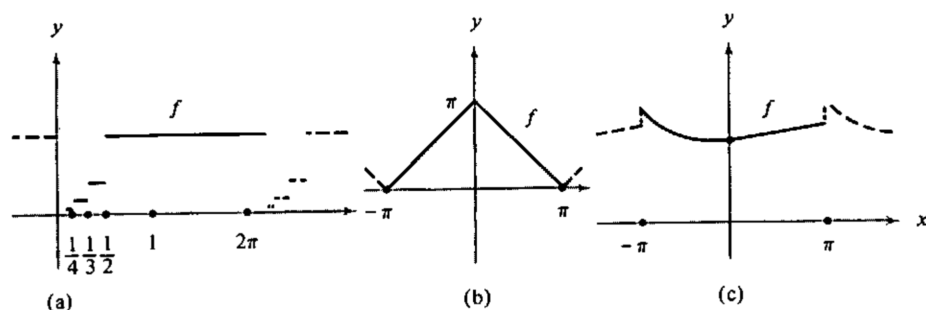


FIGURA 10.6-2 Determinación de las propiedades de convergencia de la serie de Fourier de estas funciones

La serie de Fourier en **a** converge puntualmente al punto medio del salto en una discontinuidad y a 1/2 en el origen, por el teorema de convergencia puntual.

En los casos **a** y **c** la convergencia no es uniforme, pues f no es continua (para la continuidad en los extremos hay que analizar la extensión periódica; entonces **c** desarrolla una discontinuidad).

La función en **b** tiene una serie de Fourier uniformemente convergente, pues satisface las condiciones del teorema de convergencia uniforme.

La serie de Fourier de **c** converge a f en cada x tal que $-\pi < x < \pi$ y en $-\pi$ y π converge a

$$\frac{1}{2}[f(-\pi)+f(\pi)] = \frac{1}{2}[(\pi^2+1)+(\pi+1)] = \left(\frac{\pi^2+\pi}{2}\right)+1.$$

Sólo la serie de **b** se puede derivar, para obtener la serie de Fourier de

$$f'(x) = \begin{cases} 1 & -\pi \le x \le 0 \\ -1 & 0 < x \le \pi \end{cases}$$

que converge a f' para $x \neq 0$ y a 0 en $x = 0$. ♦

Ejercicios de §10.6

Para cada uno de los ejercicios 1–3, determínese el tipo de convergencia que tiene la serie de Fourier y si podemos derivar la serie.

1. $f(x) = x^2$ en $[-\pi, \pi]$.

2. $f(x) = \pi - x^2$ en $[-\pi, \pi]$.

3. $$f(x) = \begin{cases} 3 & \text{si } -\pi \le x < -1/2 \\ 0 & \text{si } -1/2 \le x < 1/2 \\ 3 & \text{si } 1/2 \le x \le \pi. \end{cases}$$

4. Úsese el teorema **10.5.1** para determinar la serie de Fourier de x^3 en $[-\pi, \pi]$, usando la serie de Fourier de x^2 de la tabla 10.5-4.

5. **a.** Supóngase que $f: [-\pi, \pi] \to \mathbb{R}$ es derivable en $[-\pi, \pi]$, $f(-\pi) = f(\pi)$ y f' y f'' son continuas a trozos, con discontinuidades de salto. Muéstrese que

$$\frac{1}{\pi}\int_{-\pi}^{\pi} |f'(x)|^2\, dx = \sum_{n=1}^{\infty} n^2(a_n^2 + b_n^2),$$

donde a_n, b_n son los coeficientes de Fourier de f.

b. Úsese **a** y la desigualdad de Schwarz para deducir que $\sum_{n=1}^{\infty} (a_n^2 + b_n^2)^{1/2} < \infty$.

6. Considérese la serie de cosenos de sen x para un semiintervalo en $[-\pi, \pi]$. Verifíquese el teorema de derivación directamente en este caso.

§10.7 Aplicaciones

En esta sección describimos algunas aplicaciones de los métodos de Fourier a problemas de contorno simples que se presentan en la física y la ingeniería. Estos ejemplos son bastante sencillos, aunque sirven para ilustrar las técnicas básicas. Este material sólo pretende servir como ilustración y como vínculo con otros cursos de matemáticas, ingeniería o física a los que podría asistir el estudiante. De ninguna manera es un curso completo de problemas de contorno. Por ejemplo, sólo usamos las coordenadas rectangulares, cuando, de hecho, las coordenadas polares y esféricas también son muy útiles.

Los problemas que analizaremos son clásicos: la cuerda vibrante, la conducción del calor y la ecuación de Laplace. En los ejercicios 19 y 71 del final del capítulo aparecen más aplicaciones a problemas de contorno para ecuaciones diferenciales ordinarias.

Comenzaremos considerando el problema de la cuerda vibrante. Mediante argumentos físicos usuales, encontramos que el modelo matemático aproximado para una cuerda vibrante con densidad uniforme y un (pequeño) desplazamiento vertical $y(x, t)$ en x en el instante t implica que $y(x, t)$ debe satisfacer la ***ecuación de ondas***

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2}, \quad 0 \leq x \leq l. \tag{1}$$

En este caso, c es una constante determinada por la física de la cuerda y representa la velocidad de propagación de la onda a través de la cuerda (como se verá más adelante). Véase la figura 10.7-1.

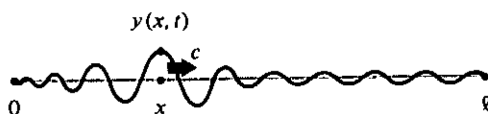


FIGURA 10.7-1 La ecuación de ondas describe la propagación de ondas en una cuerda

Para determinar completamente el problema, es necesario dar condiciones adicionales. Por ejemplo, podemos especificar las ***condiciones iniciales***, incluyendo la configuración de la cuerda en $t = 0$, esto es, cómo se "pulsa" inicialmente, o, dicho de otro modo, la función $y(x, 0)$. Físicamente, es razonable suponer que dy/dt se anula en $t = 0$ (es decir, la cuerda está inmóvil en el instante en que se pulsa). También es necesario especificar lo que ocurre en los extremos de la cuerda. Por lo general se mantienen fijos, es decir, $y(0, t) = 0$, $y(l, t) = 0$, aunque existen otras opciones. Tales especificaciones se conocen como ***condiciones de contorno.***

Una vez seleccionado este modelo de la cuerda vibrante tenemos un problema matemático; volveremos a la física cuando queramos interpretar la solución. Existe un método básico para resolver estos problemas, llamado ***separación de variables***, con el que se obtienen soluciones particulares a partir de las cuales se pueden construir soluciones generales.

Considérese el caso de un desplazamiento inicial dado. Para ser más precisos, ***el problema del desplazamiento inicial*** es el problema de determinar $y(x, t)$ para $0 < x < l$, de modo que

1. $\dfrac{\partial^2 y}{\partial t^2} = c^2 \dfrac{\partial^2 y}{\partial x^2}$ ***ecuación de movimiento***

2. en $t = 0$, $\begin{cases} y(x,0) = f(x) & \text{(para } f \text{ dada)} \\ \dfrac{\partial y}{\partial t}(x,0) = 0 & \text{(velocidad inicial nula)} \end{cases}$ ***condiciones iniciales***

3. $y(0,t) = 0,\ y(l,t) = 0$ (para todo t) ***condiciones de contorno***

Así, buscamos el movimiento de la cuerda para un tiempo futuro (o pasado) cuando ésta se "pulsa" inicialmente con la forma $f(x)$. Para que **2** y **3** sean consistentes, suponemos también que $f(0) = 0 = f(l)$. En el ejercicio 3 de esta sección consideramos otros tipos de condiciones iniciales.

La separación de variables significa que primero buscamos soluciones a la ecuación de movimiento de la forma

$$y(x,t) = h(x)g(t).$$

Al sustituir esto en la ecuación de movimiento, obtenemos

$$h(x)g''(t) = c^2 h''(x)g(t),$$

que se satisface si

$$h''(x) + \lambda h(x) = 0 \quad \text{y} \quad g''(t) + \lambda c^2 g(t) = 0 \tag{2}$$

para una constante λ (¿por qué?). Una solución de (2) con $h(0) = h(l) = 0$ y $g'(0) = 0$ es

$$h(x) = \operatorname{sen} \frac{n\pi x}{l} \quad \text{y} \quad g(t) = \cos \frac{n\pi ct}{l}$$

donde

$$\lambda = \frac{n^2\pi^2}{l^2}, \quad n = 1, 2, 3, \ldots.$$

Así, para cada n, una solución de las ecuaciones de movimiento que satisface las condiciones **1** y **3** es

$$y_n(x,t) = \operatorname{sen} \frac{n\pi x}{l} \cos \frac{n\pi ct}{l}, \quad n = 1, 2, \ldots.$$

Las condiciones iniciales de esta solución son

$$y(x,0) = \operatorname{sen}\frac{n\pi x}{l} \quad \text{y} \quad \frac{\partial y}{\partial t}(x,0) = 0.$$

Así, tenemos una solución para una condición inicial particular $\operatorname{sen}(n\pi x/l)$; sin embargo sabemos que cualquier f se puede desarrollar mediante una serie de senos en un semiintervalo, y como todas las condiciones son lineales, deberíamos poder sumar las soluciones correspondientes a los términos en este desarrollo. Esto se puede hacer rigurosamente como sigue.

10.7.1 Teorema *En el problema del desplazamiento inicial, supóngase que f es dos veces derivable. Entonces la solución a dicho problema es*

$$y(x,t) = \frac{1}{2}[f(x-ct) + f(x+ct)] = \sum_{n=1}^{\infty} b_n \operatorname{sen}\frac{n\pi x}{l}\cos\frac{n\pi ct}{l}, \tag{3}$$

donde los b_n son los coeficientes de la serie de senos en un semiintervalo,

$$b_n = \frac{2}{l}\int_0^l f(x)\operatorname{sen}\frac{n\pi x}{l}\,dx,$$

y f se extiende de modo que sea impar y periódica (dos veces derivable significa que la función f extendida es dos veces derivable; véase la figura 10.7-2).

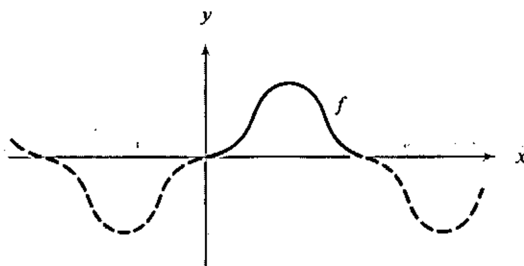


FIGURA 10.7-2 Resolución del problema del desplazamiento inicial

En este caso, la solución (3) se podría simplificar a una forma explícita más sencilla de manejar; sin embargo, con frecuencia hay que trabajar directamente con la propia serie de Fourier.

Antes de generalizarlo, observemos la sencilla interpretación física del resultado. La gráfica de $f(x - ct)$ es la de f movida hacia la derecha una distancia ct, de modo que

podemos interpretar la función $g_t(x) = f(x - ct)$ como f moviéndose hacia la derecha con velocidad c, después de un tiempo t. Análogamente, $h_t(x) = f(x + ct)$ es f moviéndose hacia la *izquierda* con velocidad c. Véase la figura 10.7-3. Así, la forma inicial de la cuerda se propaga a la izquierda y la derecha con velocidad c, formando dos ondas, cada una de las cuales tiene una amplitud igual a la mitad de la amplitud inicial, y las ondas se "reflejan" (con un cambio de signo) cuando llegan a los extremos.

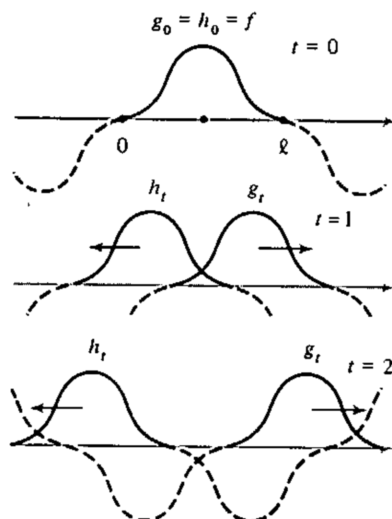


FIGURA 10.7-3 Ondas viajeras hacia la izquierda y hacia la derecha

Para usar las series de senos en semiintervalos, recordemos que hicimos que f fuera impar y periódica. Si sólo nos fijamos en el intervalo $[0, l]$, vemos que cuando f se mueve hacia l se refleja en la pared; véase la figura 10.7-4. Como la solución es la suma, habrá una cancelación complicada (o "interferencia").

Para llevar un registro de esto, es útil visualizar primero una situación más sencilla. Supóngase que f está concentrada cerca de un punto (posiblemente una función δ) y que vamos a observar su movimiento. Los movimientos en este caso reciben el nombre de ***características*** del problema. Deben visualizarse como si estuviésemos viendo una película. Véase la figura 10.7-5.

Para usar las verdaderas funciones delta o funciones f que son continuas pero no dos veces derivables, debemos generalizar el alcance del teorema **10.7.1** y también generalizar lo que entendemos por una solución de $\partial^2 y/\partial t^2 = c^2(\partial^2 y/\partial x^2)$ para y no derivable. Esto se consigue mediante la teoría de distribuciones. Si admitimos el uso de estas distribuciones, el resultado sigue siendo válido si f es una distribución (es decir,

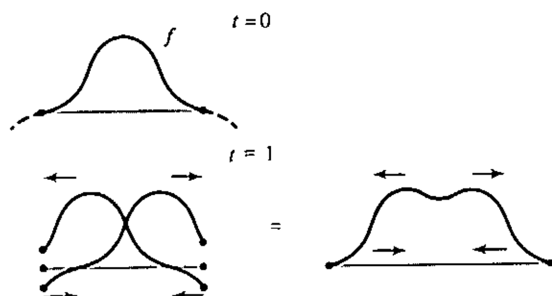


FIGURA 10.7-4 Ondas reflejadas en el contorno

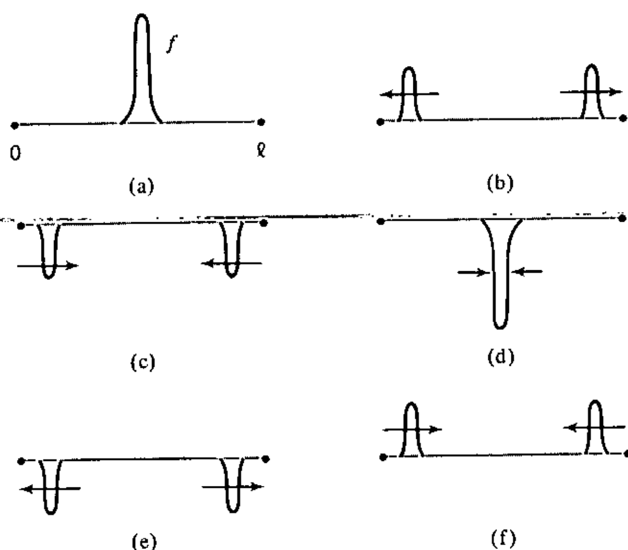


FIGURA 10.7-5 (a) $t = 0$; (b) $t = 1$; (c) $t = 2$; (d) $t = 3$; (e) $t = 4$; (f) $t = 5$; (g) regreso a (a)

las manipulaciones formales se pueden justificar si se interpretan adecuadamente). Entonces consideramos $(f(x-ct)+f(x+ct))/2$ como la solución para cualquier f, derivable o no.

En los problemas bidimensionales (como una membrana vibrante), la ecuación de ondas se lee

$$\frac{\partial^2 y}{\partial t^2} = c^2 \left(\frac{\partial^2 y}{\partial x_1^2} + \frac{\partial^2 y}{\partial x_2^2} \right) \tag{4}$$

En este caso la solución general se puede escribir como una serie de Fourier, pero no tiene una expresión explícita sencilla como en el caso unidimensional. La solución

(para el problema similar del desplazamiento inicial) en el rectángulo $[0, l] \times [0, l']$ está dada por

$$y(x_1, x_2, t) = \sum_{n,m=1}^{\infty} b_{nm} \operatorname{sen}\frac{n\pi x_1}{l} \operatorname{sen}\frac{m\pi x_2}{l'} \cos\left[\pi ct\sqrt{(n/l)^2 + (m/l')^2}\right] \tag{5}$$

donde

$$b_{nm} = \frac{4}{ll'}\int_0^l \int_0^{l'} f(x_1, x_2) \operatorname{sen}\frac{n\pi x_1}{l} \operatorname{sen}\frac{m\pi x_2}{l'}\, dx_1\, dx_2.$$

Se pedirá al lector que obtenga esta expresión en el ejercicio 68 del final de este capítulo.

Pasemos ahora a la conducción del calor. Considérese una barra cuya temperatura es $T(x, t)$ en el punto x en el instante t. Si interpretamos $-(\partial T/\partial x)$ como la tasa de flujo de calor, la condición de "aislamiento" en $x = 0$ es $\partial T/\partial x = 0$ (evaluada en $x = 0$). La ley de conducción del calor postula que

$$\frac{\partial T}{\partial t}(x, t) = k\frac{\partial^2 T}{\partial x^2}(x, t), \tag{6}$$

donde k es una constante determinada por la conductividad del material. Ésta es la ***ecuación del calor*** (para ver una forma de obtenerla, consúltese Marsden y Tromba, *Cálculo vectorial*, *op.cit.*, capítulo 7).

La ecuación del calor (6) difiere de la ecuación de ondas en que $\partial T/\partial t$ reemplaza a $\partial^2 T/\partial t^2$. Esta diferencia es muy importante, ya que las soluciones del problema de conducción del calor tienen un comportamiento diferente al de la ecuación de ondas. Por ejemplo, en la ecuación del calor sólo se obtienen soluciones para $t \geq 0$. Intuitivamente, la gráfica de la solución de la ecuación de ondas "rebota" continuamente como si fuera una onda en el agua. Para la ecuación del calor, la solución se difunde y se vuelve estacionaria cuando $t \to \infty$ (ya que la temperatura tiende a ser uniforme).

Para estudiar esta sencilla situación crearemos el siguiente modelo para la conducción del calor de una barra con extremos aislados (para simplificar, sea $k = 1$). En consecuencia, queremos determinar una $T(x, t)$ que satisface

1. $\dfrac{\partial T}{\partial x}(x, t) = k\dfrac{\partial^2 T}{\partial x^2}(x, t), \quad 0 < x < l,\ t \geq 0$ — ***ecuación del calor***

2. $T(x, 0) = f(x), \quad 0 < x < l$ — ***condiciones iniciales***

y

3. $\begin{cases} \dfrac{\partial T}{\partial x}(0, t) = 0 \\ \dfrac{\partial T}{\partial x}(l, t) = 0 \end{cases} \quad t \geq 0$ — ***condiciones de contorno***

En primer lugar, determinamos soluciones particulares para una f particular mediante separación de variables. Si $T(x, t) = g(x)h(t)$, obtenemos

$$g(x)h'(t) = g''(x)h(t).$$

(recuérdese que $k = 1$). Estas ecuaciones son ciertas si, para una constante λ,

$$g(x) + \lambda g''(x) = 0 \quad \text{y} \quad h(t) + \lambda h'(t) = 0. \tag{7}$$

Las soluciones de estas ecuaciones que satisfacen las condiciones de contorno son

$$g(x) = \cos\frac{n\pi x}{l} \quad \text{y} \quad h(t) = e^{-n^2\pi^2 t/l^2}, \quad n = 0, 1, 2, \ldots,$$

donde $\lambda = n^2\pi^2/l^2$. Usamos el coseno y no el seno de modo que se cumpla la tercera condición de contorno (estas condiciones de contorno no se pueden cumplir si intentamos determinar g y h para $\lambda < 0$). Así, una solución con $f(x) = \cos(n\pi x/l)$ está dada por $e^{-n^2\pi^2 t/l^2}\cos(n\pi x/l)$, $n = 0, 1, 2, \ldots$. Como todas las expresiones son lineales y

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\frac{n\pi x}{l}$$

(serie de cosenos en un semiintervalo), esperamos que la solución general con condición inicial f esté dada por

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n e^{-n^2\pi^2 t/l^2} \cos\frac{n\pi x}{l}.$$

10.7.2 Teorema *Si f es de cuadrado integrable, entonces, para cada $t > 0$,*

$$T(x,t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n e^{-n^2\pi^2 t/l^2} \cos\frac{n\pi x}{l} \tag{8}$$

converge uniformemente, es derivable y satisface la ecuación del calor y las condiciones de contorno. En $t = 0$ es igual a f en sentido de la convergencia en media, y puntualmente si f es de clase C^1. Como siempre,

$$a_n = \frac{2}{l}\int_0^l f(x)\cos\frac{n\pi x}{l}\,dx.$$

El término exponencial en (8) hace que la convergencia sea rápida para $t > 0$. Para $t < 0$, por lo general hay divergencia. Cuando $t \to \infty$ todos los términos de la serie convergen a 0 y también la suma converge a 0, de modo que

$$\lim_{t\to\infty} T(x,t) = \frac{1}{2}a_0,$$

y entonces T se convierte en una temperatura constante uniforme de acuerdo con nuestra intuición. Pediremos al lector que demuestre esto en el ejercicio 69 al final del capítulo.

El siguiente resultado, más fino, describe lo que ocurre si $t \to 0$.

10.7.3 Teorema *En el teorema* 10.7.2,

$$\lim_{t\to 0,\ t>0} T(x,t) = f(x) \tag{9}$$

en el sentido de la convergencia en media, y la convergencia es uniforme (y puntual) si f es continua, con f' continua a trozos. De modo más general, si la serie de Fourier de f converge en x a $f(x)$, entonces $T(x, t) \to f(x)$ cuando $t \to 0$.

Este importante resultado nos dice el sentido en que recuperamos la condición inicial a partir de los valores $t > 0$. Esto no es consecuencia de la derivabilidad de $T(x, t)$ para $t > 0$.

Nuestra última aplicación será el estudio de la ecuación de Laplace en un cuadrado. La ***ecuación de Laplace*** en $\mathbb{R}^n$ es

$$\nabla^2\varphi = 0;\quad \text{es decir,}\quad \sum_{k=1}^{n} \frac{\partial^2\varphi}{\partial x_k^2} = 0. \tag{10}$$

Tal función φ se denomina ***armónica***. Esta ecuación surge en muchos problemas de electrostática, flujo de fluidos y conducción del calor. Por ejemplo, en la conducción del calor en $\mathbb{R}^n$, si T es independiente de t, de modo que $\partial T/\partial t = 0$, entonces T es solución de la ecuación de Laplace.

El problema básico, llamado ***problema de Dirichlet***, es el siguiente: Dados los valores de φ en alguna curva cerrada en el plano, encontrar φ en el interior. Este problema aparentemente sencillo es el núcleo del vasto tema de la teoría de potencial (la terminología surge de la electrostática, donde φ representa el potencial eléctrico; de nuevo los detalles se pueden ver en Marsden y Tromba, *Cálculo vectorial*, capítulo 7; este problema también se puede atacar mediante métodos de variable compleja; véase, por ejemplo, J. Marsden y M. Hoffman, *Basic Complex Analysis*, *op.cit.*, capítulo 5).

Usaremos las series de Fourier para resolver este problema para un cuadrado en $\mathbb{R}^2$ (para cubos en $\mathbb{R}^3$ los resultados son similares). El problema se resume como sigue.

En $[0, a] \times [0, b] \subset \mathbb{R}^2$, determínese una función φ tal que

1. $\nabla^2\varphi = 0$ ***ecuación de Laplace***

2. $\varphi(x,0) = g_1(x), \quad \varphi(x,b) = g_2(x),$
$\varphi(0,y) = f_1(y), \quad \varphi(a,y) = f_2(y),$ ***condiciones de contorno***

donde f_i y g_j son funciones dadas. Véase la figura 10.7-6.

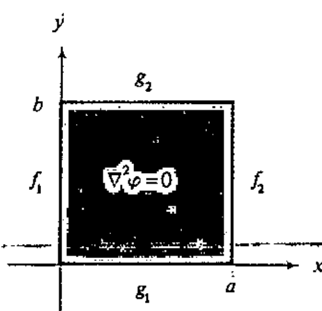


FIGURA 10.7-6 Resolución de la ecuación de Laplace con condiciones de contorno dadas

Primero obtendremos unas soluciones particulares mediante separación de variables. Si $\varphi(x, y) = \varphi_1(x)\varphi_2(y)$, obtenemos

$$\varphi_1''(x)\varphi_2(y) + \varphi_1(x)\varphi_2''(y) = 0,$$

que es válida si, para una constante λ, las funciones φ_1 y φ_2 satisfacen las ecuaciones

$$\varphi_1''(x) + \lambda\varphi_1(x) = 0 \quad \text{y} \quad \varphi_2''(y) - \lambda\varphi_2(y) = 0. \tag{11}$$

Las soluciones son

$$\varphi_1(x) = \operatorname{sen}\frac{n\pi x}{a}, \quad n = 1, 2, \ldots,$$

y

$$\varphi_2(y) = \operatorname{senh}\frac{n\pi(b - y)}{a},$$

donde $\lambda = n^2\pi^2/a^2$. Elegimos $\mathrm{senh}(z) = (e^z - e^{-z})/2$, $z = n\pi(b - y)/a$ en vez de e^z o e^{-z}, pues éstas se anulan cuando $y = b$. Análogamente, elegimos el seno en vez del coseno. Así,

$$\varphi(x, y) = \mathrm{senh}\,\frac{n\pi(b - y)}{a}\,\mathrm{sen}\frac{n\pi x}{a}$$

satisface las condiciones de contorno

$$g_1 = \mathrm{senh}\,\frac{n\pi b}{a}\,\mathrm{sen}\frac{n\pi x}{a}, \quad f_1 = f_2 = g_2 = 0.$$

Podemos obtener otras soluciones básicas de forma similar. Por lo tanto, es de esperar que cuando $f_1 = f_2 = g_2 = 0$, la solución del problema sea

$$\varphi(x, y) = \sum_{n=1}^{\infty} b_n \mathrm{senh}\,\frac{n\pi(b - y)}{a}\,\frac{\mathrm{sen}(n\pi x/a)}{\mathrm{senh}(n\pi b/a)}, \tag{12}$$

donde $g_1(x) \sum_{n=1}^{\infty} b_n\, \mathrm{sen}(n\,\pi\, x/a)$ (serie de senos en un semiintervalo). Existen soluciones similares para los otros lados, y la suma es la solución para todos los lados.

10.7.4 Teorema

i. *Dado g_1, sea $\varphi(x, y)$ dada por (12). Supóngase que g_1 es de clase C^2 y $g_1(0) = g_1(a) = 0$. Entonces φ converge uniformemente, es la solución al problema de Dirichlet con $f_1 = f_2 = g_2 = 0$, y es continua en todo el cuadrado y $\nabla^2\varphi = 0$ en el interior.*

ii. *Si cada f_1, f_2, g_1, g_2 es de clase C^2 y se anula en las esquinas del rectángulo, entonces la solución $\varphi(x, y)$ es la suma de cuatro series como la de la ecuación (12), $\nabla^2\varphi = 0$ en el interior, φ es continua en todo el rectángulo y alcanza los valores dados sobre el contorno. Además, φ es C^∞ en el interior.*

iii. *Si f_1, f_2, g_1, g_2 sólo son de cuadrado integrable, entonces la serie de φ converge en el interior, $\nabla^2\varphi = 0$ y φ es C^∞. Además, φ alcanza los valores de contorno en el sentido de la convergencia en media. Esto significa, por ejemplo, que* $\mathrm{lím}_{y\to 0}\, \varphi(x, y) = \varphi(x, 0) = g_1(x)$ *con la convergencia en media.*

Los resultados **i** y **ii** siguen siendo válidos si sólo suponemos que f_i y g_i son continuas, pero requieren un método distinto de demostración. Este procedimiento es bueno, pues da la solución explícita en términos de series de Fourier.

10.7.5 Ejemplo *En el problema del desplazamiento inicial, defínase la **energía total** de la cuerda en el instante t como*

$$E(t) = \frac{1}{2}\int_0^l \left(\frac{\partial y}{\partial t}\right)^2 dx + \frac{c^2}{2}\int_0^l \left(\frac{\partial y}{\partial x}\right)^2 dx \tag{13}$$

(energía cinética más energía potencial). Muéstrese que E(t) es constante con respecto a t.

Solución Basta mostrar que $dE/dt = 0$. Ahora,

$$\frac{dE}{dt} = \frac{1}{2}\int_0^l \frac{\partial}{\partial t}\left(\frac{\partial y}{\partial t}\right)^2 dx + \frac{c^2}{2}\int_0^l \frac{\partial}{\partial t}\left(\frac{\partial y}{\partial x}\right)^2 dx. \tag{14}$$

Esto se justifica mediante la derivación bajo el signo de la integral si y es continuamente derivable dos veces (véase el capítulo 9). Entonces

$$\frac{c^2}{2}\int_0^l \frac{\partial}{\partial t}\left(\frac{\partial y}{\partial x}\right)^2 dx = c^2\int_0^l \frac{\partial y}{\partial x}\frac{\partial^2 y}{\partial t\,\partial x}\,dx. \tag{15}$$

Al integrar el miembro derecho de (15) por partes y usar el hecho de que $\partial y/\partial t$ en $x = 0$, l se obtiene

$$-c^2\int_0^l \frac{\partial y}{\partial t}\frac{\partial^2 y}{\partial x^2}\,dx,$$

lo que, en vista de la ecuación de movimiento, es igual a

$$-\int_0^l \frac{\partial y}{\partial t}\frac{\partial^2 y}{\partial t^2}dx.$$

Así, $dE/dt = 0$, pues el primer término de (14) es

$$\int_0^l \frac{\partial y}{\partial t}\frac{\partial^2 y}{\partial t^2}\,dx. \quad \blacklozenge$$

Si y no es dos veces derivable hay que tener más cuidado (por ejemplo, si y es una función δ, E ni siquiera está definida). Para la ecuación del calor no tenemos conservación de la energía, pues, intuitivamente, la energía se difunde (véase el ejercicio 6 de esta sección).

10.7.6 Ejemplo *Una barra con extremos aislados tiene una distribución de temperatura dada por* $f(x) = x$, $0 < x < l$ en $t = 0$. *Determínese la distribución de temperatura para* $t > 0$.

Solución De acuerdo con el teorema **10.7.2**, usamos la serie de cosenos en semi-intervalos para x e insertamos factores $e^{-n^2\pi^2 t/l^2})^2$. La serie de x está dada por

$$x = \frac{l}{2} - \frac{4l}{\pi^2}\sum_{n=1}^{\infty}\frac{\cos[(2n-1)\pi x/l]}{(2n-1)^2},$$

de modo que la solución pedida es

$$T(x,t) = \frac{l}{2} - \frac{4l}{\pi^2}\sum_{n=1}^{\infty} e^{-((2n-1)^2\pi^2 t)/l^2}\frac{\cos[(2n-1)\pi x/l]}{(2n-1)^2}.$$

En general no se puede reducir esta expresión a una forma más compacta, sino que debemos trabajar con este desarrollo en serie. ♦

10.7.7 Ejemplo *Resuélvase el problema de Dirichlet en* $[0,\pi]\times[0,\pi]$ *con valores de contorno* $g_1 = 1$, $f_1 = 0$, $f_2 = 0$, $g_2 = 0$. *¿Cómo se alcanzan los valores de contorno?*

Solución La serie de senos para 1 es

$$\frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\operatorname{sen}[(2n-1)x]}{2n-1}.$$

De acuerdo con el teorema **10.7.4**, φ se obtiene al insertar factores

$$\frac{\operatorname{senh}[n(\pi-y)]}{\operatorname{senh}(n\pi)},$$

de modo que obtenemos

$$\varphi(x,y) = \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\operatorname{senh}[(2n-1)(\pi-y)]}{\operatorname{senh}[(2n-1)\pi]}\frac{\operatorname{sen}[(2n-1)x]}{2n-1}$$

como solución. Por el teorema **10.7.4**, φ es C^∞ y satisface $\nabla^2\varphi = 0$ en el interior del cuadrado. En este caso, $\varphi(x, 0) = 1$ en el sentido de la convergencia en media; es decir, $\varphi(x, y) \to 1$ cuando $y \to 0$ (en media). La suma parcial s_n de φ se bosqueja en la figura 10.7-7.

Ejercicios de §10.7

1. Para el problema de desplazamiento inicial de una cuerda, considérese la "cuerda pulsada" tal que $f(x) = hx$ si $0 \le x \le l/2$ y $f(x) = lh - hx$ si $l/2 \le x \le l$. Muéstrese

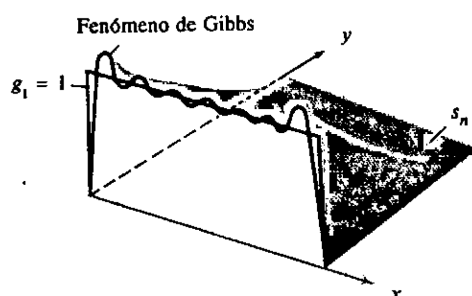


FIGURA 10.7-7 Una solución específica del problema de Dirichlet

que f no satisface las hipótesis del teorema **10.7.1**. Encuéntrese una expresión para la solución después del instante $t = l/c,\ 3l/2c$.

2. Supóngase que, en el problema del desplazamiento inicial, f tiene un máximo o una discontinuidad en x_0. Muéstrese que, después del instante t, esta propiedad se propaga como una característica.

3. **a.** Enúnciese el problema del desplazamiento inicial para una cuerda sin desplazamiento inicial. Si la velocidad inicial es $g(x)$, muéstrese que la solución es

$$y(x,t) = \frac{1}{2c}\left\{\int_0^{x+ct} g(z)\,dz - \int_0^{x-ct} g(z)\,dz\right\}.$$

   b. Combínese **a** con el problema del desplazamiento inicial para obtener una solución del problema con desplazamiento y velocidad iniciales (ésta es la ***solución de d'Alembert***).

4. En el teorema **10.7.2**, demuéstrese que para cada $t \geq 0$, $(2/l)\int_0^l T(x,t)dx = a_0$.

5. Una barra con extremos aislados tiene en $t = 0$ la distribución de temperaturas $f(x, 0) = x^2$. Determínese la temperatura para $t > 0$ y el límite cuando $t \to \infty$.

6. Sea $T(x, t)$ una solución de la ecuación del calor y sea $L(t) = \int_0^l |T(x,t)|^2 dx$. Muéstrese que $L(t)$ es no creciente.

§10.8 Integrales de Fourier

Esta sección contiene un estudio informal de las integrales de Fourier. Bosquejaremos los resultados principales de modo que el lector pueda ver el papel que desempeñan en el análisis de Fourier. Como hemos visto en las secciones anteriores, las series de Fourier

son una útil herramienta para el análisis de funciones en un intervalo finito. Como muchas funciones están definidas en toda la recta real $\mathbb{R}$, sería útil tener una teoría similar en $\mathbb{R}$. Las integrales de Fourier nos proporcionan esta teoría.

Considérese $f : [-l, l] \to \mathbb{R}$. Si escribimos f en términos de su serie de Fourier exponencial,

$$f(x) = \sum_{-\infty}^{\infty} c_n e^{in\pi x/l}, \qquad \text{donde } c_n = \frac{1}{2l}\int_{-l}^{l} f(y)e^{-in\pi y/l}\,dy.$$

Sea $\alpha = n\pi/l$ y escribamos

$$f(x) = \sum_{-\infty}^{\infty} c(\alpha)e^{i\alpha x}\frac{\pi}{l}, \qquad \text{donde } c(\alpha) = \frac{1}{2\pi}\int_{-l}^{l} f(y)e^{-i\alpha y}\,dy.$$

Para l grande, α se aproxima a una variable continua y esta suma es una suma de Riemann con $\Delta\alpha = \pi/l$. Esto sugiere que

$$f(x) = \int_{-\infty}^{\infty} c(\alpha)e^{i\alpha x}\,d\alpha, \qquad \text{donde } c(\alpha) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(y)e^{-i\alpha y}\,dy. \tag{1}$$

En resumen, podemos extender nuestros intervalos finitos a intervalos infinitos, y la serie de Fourier se expresa entonces como una integral.

Podemos dar los mismos pasos en forma trigonométrica, excepto que las integrales se toman de 0 a ∞, al igual que las sumas correspondientes.

El teorema importante en este contexto establece que si f es continua a trozos con discontinuidades de salto, $f'(x_0^+)$ y $f'(x_0^-)$ existen en esos puntos y $\int_{-\infty}^{\infty}|f(x)|dx < \infty$ (f es integrable), por lo que

$$\frac{1}{2}[f(x^+) + f(x^-)] = \int_{-\infty}^{\infty} c(\alpha)e^{i\alpha x}\,d\alpha, \tag{2}$$

donde

$$c(\alpha) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(y)e^{-i\alpha y}\,dy.$$

Esto se demuestra del mismo modo que el teorema de la convergencia puntual.

La fórmula (2) es la ***fórmula de la transformada inversa de Fourier***. En forma trigonométrica, la fórmula es

$$\frac{1}{2}[f(x^+) + f(x^-)] = \int_{0}^{\infty} [A(\alpha)\cos\alpha x + B(\alpha)\operatorname{sen}\alpha x]\,d\alpha, \tag{3}$$

donde

$$A(\alpha) = \frac{1}{\pi}\int_0^{\infty} f(y)\cos\alpha y\,dy \qquad \text{y} \qquad B(\alpha) = \frac{1}{\pi}\int_0^{\infty} f(y)\operatorname{sen}\alpha y\,dy.$$

Esta forma es de particular conveniencia si f es par o impar.

En vista del teorema de la transformada inversa, la ***transformada de Fourier*** de f se define como

$$\hat{f}(\alpha) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-ix\alpha}\,dx. \tag{4}$$

Si f es continua, derivable e integrable en $\mathbb{R}$, entonces

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\alpha)e^{i\alpha x}\,d\alpha. \tag{5}$$

Existe una fórmula análoga en $\mathbb{R}^n$; es decir,

$$f(x) = \int_{\mathbb{R}^n} \hat{f}(\alpha)e^{i\langle\alpha,\,x\rangle}\,d\alpha, \tag{6}$$

donde

$$\hat{f}(\alpha) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} f(x)e^{-i\langle x,\alpha\rangle}\,dx,$$

$x, \alpha \in \mathbb{R}^n$ y $\langle x, \alpha\rangle$ es el producto escalar usual en $\mathbb{R}^n$.

Dada $f: [0, \infty[\to \mathbb{R}$, podemos extenderla a todo $\mathbb{R}$ haciéndola par o impar. Al igual que con las series de senos y de cosenos, podemos utilizar la ***transformada coseno de Fourier*** extendiendo f para que sea par, y haciendo

$$\hat{f}_c(\alpha) = \frac{2}{\pi}\int_0^{\infty} f(y)\cos(\alpha y)\,dy. \tag{7}$$

La fórmula de la transformada inversa resulta ser

$$f(x) = \int_0^{\infty} \hat{f}_c(\alpha)\cos(\alpha x)\,d\alpha. \tag{8}$$

Análogamente, extender f para que sea impar nos conduce a la ***transformada seno de Fourier***,

$$\hat{f}_s(\alpha) = \frac{2}{\pi}\int_0^{\infty} f(y)\mathrm{sen}(\alpha y)\,dy, \tag{9}$$

y la fórmula de la transformada inversa es

$$f(x) = \int_0^{\infty} \hat{f}_s(\alpha)\mathrm{sen}(x\alpha)\,d\alpha. \tag{10}$$

Un hecho conocido que debe aprenderse (usando el ejemplo resuelto **9.1**, capítulo 9) es que la transformada de Fourier de $e^{-x^2/2}$ (la función gaussiana) en $\mathbb{R}$ es $e^{-\alpha^2/2}/\sqrt{2\pi}$, lo que es consistente con el teorema de la transformada inversa.

En general, una ***transformada integral*** es una asociación de la función

$$g(x) = \int_A k(x, y) f(y)\, dy \tag{11}$$

con la función f para alguna función fija k llamada ***núcleo*** y algún rango fijo A de integración. Tales operaciones son comunes en la física matemática. Así, la transformada de Fourier es una transformada integral con núcleo $k(x, y) = e^{-ixy}/2\pi$. Aquí llegamos a un importante problema general. La transformada aplica f en g; es decir, dada f, obtenemos g. ¿Podemos invertir esto? En otras palabras, dada g, ¿podemos invertir la transformación para determinar f?

La fórmula de la transformada inversa de Fourier resuelve este problema en el caso $k(x, y) = e^{-ixy}/2\pi$. Es decir, si conocemos la transformada de Fourier, podemos recupe-la función por medio de la fórmula de la transformada inversa.

Otra transformada integral común es la ***transformada de Laplace,*** con núcleo $k(x, y) = e^{-xy}$ y rango $[0, \infty]$. Así, la transformada de Laplace de f es

$$\mathcal{L}(f)(x) = \int_0^\infty e^{-xy} f(y)\, dy. \tag{12}$$

El problema de inversión para las transformadas de Laplace también tiene una solución (véanse los detalles, por ejemplo, en Marsden y Hoffman, *Basic Complex Analysis, op. cit.*, capítulo 7).

Para $f : [-\pi, \pi] \to \mathbb{R}$, la relación de Parseval produce la identidad (véase la tabla 10.5-3)

$$\|f\|^2 = \int_{-\pi}^{\pi} |f(x)|^2\, dx = 2\pi \sum_{-\infty}^{\infty} |c_n|^2 \tag{13}$$

para los coeficientes de Fourier c_n en la forma exponencial. Como c_n es análogo a las transformadas de Fourier, es de esperar que algo similar ocurra en términos de $\hat{f}$. Esto es cierto si $|f|^2$ y $|f|$ son integrables; así si

$$\|f\|^2 = \int_{-\infty}^{\infty} |f(x)|^2\, dx,$$

tenemos

$$\|f\|^2 = 2\pi \|\hat{f}\|^2. \tag{14}$$

De modo más general, $\langle f, g \rangle = 2\pi \langle \hat{f}, \hat{g} \rangle$. En este caso, la transformada de Fourier puede ser de cualquier tipo: exponencial, trigonométrica, de senos o de cosenos.

La fórmula (14) se conoce como ***relación de Parseval*** y ***teorema de Plancherel.*** Los detalles técnicos relativos a este tema requieren la teoría de integración de Lebesgue.

Si f y g son integrables en $\mathbb{R}$ (o en $\mathbb{R}^n$), definimos la ***convolución*** de f y g como

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x-y)g(y)\,dy. \tag{15}$$

Esta operación aparece de forma natural en muchos problemas. Una de sus principales propiedades es que

$$(f * g)\hat{} = 2\pi\hat{f}\cdot\hat{g} \tag{16}$$

(véase el ejemplo resuelto **10.4** al final de este capítulo).

Las transformadas de Fourier tienen aplicaciones importantes en matemática pura y aplicada. Por ejemplo, son importantes en las ecuaciones diferenciales en derivadas parciales, como la ecuación de ondas en todo $\mathbb{R}^n$. La razón es que, en términos de las transformadas de Fourier, las ecuaciones se vuelven más sencillas, a menudo algebraicas, y cuando se resuelven estas ecuaciones, la respuesta final se obtiene mediante la fórmula de la transformada inversa. Las convoluciones aparecen en la inversión.

Mediante las transformadas de Fourier, muchos de los problemas resueltos en secciones anteriores sobre intervalos finitos se pueden traducir fácilmente a problemas en toda la recta real $\mathbb{R}$. Los siguientes ejercicios dan un esquema de la forma de hacerlo.

Ejercicios de §10.8

Estos problemas se pueden resolver de manera informal, ignorando el rigor, de la misma forma en que hemos introducido esta sección.

1. Muéstrese que si $\hat{f}(\alpha)$ es la transformada de Fourier de f, entonces $\hat{f}(\alpha) = i\alpha\hat{f}(\alpha)$ y $2\pi\int_{-\infty}^{\infty}\alpha^2|\hat{f}|^2\,d\alpha = \int_{-\infty}^{\infty}|f'|^2\,dx$.

2. Sea $f(x, y)$ una función que satisface $\partial^2 f/\partial x^2 + \partial^2 f/\partial y^2 = 0$. Supóngase que $f(x, 0) = g(x)$ y $\text{lím}_{y\to\infty} f(x, y)$ para todo x.

a. Sea $\hat{f}(\alpha, y)$ la transformada de Fourier de $f(x, y)$ con y considerado como constante. Muéstrese que $\hat{f}(\alpha, y) = \hat{g}(\alpha)e^{-|\alpha|y}$.

b. Muéstrese que $e^{-|\alpha|y}$ es la transformada de Fourier, con respecto a x, de

$$\frac{2y}{x^2+y^2}.$$

c. Dedúzcase que la solución de la ecuación de Laplace es

$$f(x, y) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{y}{y^2+(x-x')^2}g(x')\,dx'$$

3. Supóngase que $f(x, t)$ es una función para la cual

$$\frac{\partial f}{\partial t} = k^2 \frac{\partial^2 f}{\partial x^2}, \quad -\infty < x < \infty, \quad t \geq 0,$$

y $f(x, 0) = g(x)$. Sea $\hat{f}(\alpha, t)$ la transformada de Fourier de f.

a. Muéstrese que $\hat{f}(\alpha, t) = \hat{g}(\alpha)e^{-k^2\alpha^2 t}$.

b. Dedúzcase que la solución de la ecuación del calor es

$$f(x,t) = \frac{1}{2k\sqrt{\pi t}} \int_{-\infty}^{\infty} g(x')e^{-(x-x')^2/4k^2 t}\, dx'.$$

4. Supóngase que $f(x,t)$ es una función para la que

$$\frac{\partial^2 f}{\partial t^2} = \frac{\partial^2 f}{\partial x^2}, \quad f(x,0) = g(x), \quad \text{y} \quad \frac{\partial f}{\partial t}(x,0) = h(x).$$

Sea $\hat{f}(\alpha, t)$ la transformada de Fourier de $f(x, t)$, considerando t como constante.

a. Muéstrese que $\hat{f}(\alpha, t) = \hat{g}(\alpha)\cos \alpha t + h(\alpha)([\text{sen } \alpha t]/\alpha)$.

b. Dedúzcase que la solución de la ecuación de onda es

$$f(x,t) = \frac{1}{2}[g(x - t) + g(x + t)] + \frac{1}{2} \int_0^t [h(x - \tau) + h(x + \tau)]\, d\tau.$$

§10.9 Formalismo de la mecánica cuántica

Esta sección explicará la forma en que la teoría de series de Fourier en un espacio con producto interno (desarrollada en **§10.1** y **§10.2**) se relaciona con el formalismo de la mecánica cuántica.

En primer lugar, explicaremos algunas diferencias entre la mecánica clásica y la mecánica cuántica. En la mecánica clásica, el movimiento de una partícula se describe mediante una trayectoria definida, con su vector velocidad asociado. En mecánica cuántica siempre existe cierta "incertidumbre" en la posición o la velocidad (o en ambas). Para los fenómenos atómicos, esta incertidumbre es necesaria, de modo que tales efectos están fuera del dominio de aplicabilidad de la mecánica clásica. Por ejemplo, si se prepara una partícula atómica con una velocidad inicial definida y se proyecta sobre una pantalla, no sabremos con exactitud la trayectoria que tendrá, sino que al buscar la partícula sólo podremos determinar la probabilidad de encontrarla en una región dada.

En la figura 10.9-1 se consideran las partículas que se proyectan sobre una placa de detección, pasando a través de una pantalla con dos rendijas.[3] Sólo podemos determinar la probabilidad de la posición, no la posición exacta. Tras una serie de experimentos sucesivos y repetidos se obtienen áreas iluminadas y áreas oscuras sobre la pantalla, correspondientes a zonas de alta o baja probabilidad. Esto se representa mediante una curva en la figura. Otros fenómenos físicos, como los espectros discretos de los átomos, también requieren una descripción en términos de mecánica cuántica, como lo explican los textos sobre dicha materia.

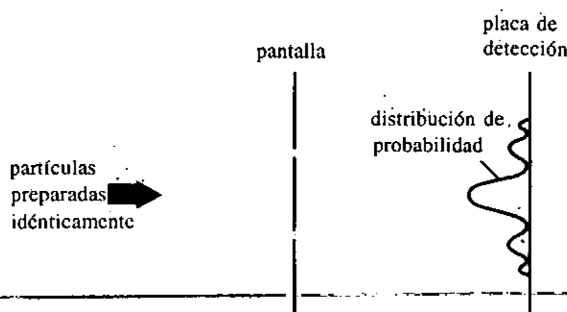


FIGURA 10.9-1 El experimento de la doble rendija.

Esta información no debe entenderse en el sentido de que la mecánica clásica sea "falsa" y que la mecánica cuántica sea "verdadera". Ambos son modelos matemáticos que aproximan bien la naturaleza dentro de sus propios rangos de aplicabilidad. La mecánica cuántica, por ejemplo, no es muy útil cuando se estudia la dinámica del transbordador espacial.

La cuestión es, entonces, ¿cómo describir el comportamiento de una partícula en mecánica cuántica? Una sola partícula en mecánica cuántica es descrita por una función de valores complejos, $\psi(x)$, donde $x \in \mathbb{R}^3$. Para más partículas, se cambia $\mathbb{R}^3$ por otro espacio (para N partículas, se usa $\mathbb{R}^{3N}$). Si el sistema depende del tiempo, entonces usamos $\psi(t, x)$. La densidad de probabilidad para la posición de la partícula en el espacio está dada por $\psi(x)\overline{\psi(x)}$, de modo que si la probabilidad total es 1, entonces deberíamos tener

$$\|\psi\|^2 = \int_{\mathbb{R}^3} \psi(x)\overline{\psi(x)}\, dx = 1,$$

es decir, que ψ esté normalizada. Esta última afirmación proporciona uno de los vínculos entre el estudio del modelo matemático (es decir, el estudio del espacio con produc-

[3] Éste es un experimento imaginario que se usa solamente a modo de ilustración. En los experimentos reales, la "pantalla" podría ser, por ejemplo, un cristal, y las ranuras podrían corresponder a los espacios interatómicos.

to interno de funciones de cuadrado integrables ψ, y las operaciones definidas en él, analizadas en esta sección) y la interpretación física de este modelo.

Si estamos midiendo una cantidad definida, como la coordenada x, el momento o el momento angular, entonces ésta no se puede medir con total certidumbre. El aspecto de la mecánica cuántica que queremos explorar es la estructura matemática de las funciones de onda ψ y los objetos matemáticos correspondientes a las cantidades físicamente medibles. Por supuesto, el tema es mucho más profundo que esto y nuestro análisis solamente roza la superficie.

Introduciremos algo de rigor en nuestro análisis. En primer lugar, supóngase que $V = L^2$ es el espacio de funciones $\psi : \mathbb{R}^3 \to \mathbb{C}$ de cuadrado integrable:

$$||\psi||^2 = \langle \psi, \psi \rangle = \int \psi(x)\overline{\psi(x)}\, dx < \infty.$$

Como hemos visto antes, éste es un espacio con producto interno cuyo producto escalar es

$$\langle f, g \rangle = \int_{\mathbb{R}^3} f(x)\overline{g(x)}\, dx.$$

Así, todo el análisis relativo a un espacio con producto interno (véase **§10.1** y **§10.2**) se puede aplicar a este caso.

Un ***estado cuántico*** es (por definición) una función $\psi \in V$ tal que $||\psi|| = 1$ (es decir, ψ está normalizada). Un ***observable*** es un operador A sobre V que es ***simétrico*** (o autoadjunto, o hermítico); esto significa que $A : V \to V$ es una transformación lineal que satisface

$$\langle Af, g \rangle = \langle f, Ag \rangle$$

para todo $f, g \in V$. En realidad, A podría estar definido solamente sobre algunos elementos de V. Por ejemplo, si A es el laplaciano, es decir, si

$$Af = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = \nabla^2 f,$$

entonces A está definido sobre aquellas $f \in V$ cuyas segundas derivadas también pertenecen a V. Se puede verificar formalmente que ∇^2 es simétrico, usando la integración por partes dos veces. Por otro lado, $\partial/\partial x$ no es simétrico pero $i(\partial/\partial x)$ sí lo es (el lector debe demostrar esta afirmación). (Los operadores como el laplaciano se denominan ***no acotados***, y para ellos hay que distinguir entre simétricos y autoadjuntos. Para mayor información, véase Reed-Simon, *Methods of Modern Mathematical Physics, vol. I, Functional Analysis*, Academic Press, Nueva York, 1972.)

Una ***autofunción*** de un operador A es una función $f \in V$, $f \neq 0$ tal que $Af = \lambda f$ para algún número complejo λ llamado ***autovalor***. Obsérvese que si f es una autofunción, entonces también lo es $f/||f||$, por lo que podemos suponer que nuestras autofunciones están normalizadas.

Debemos hacer dos observaciones importantes en relación con las autofunciones de los operadores simétricos. En primer lugar, *si A es simétrico y f es una autofunción con autovalor λ, entonces λ es real.*

Demostración $\langle Af, f\rangle = \langle \lambda f, f\rangle = \lambda \|f\|^2$. Por otro lado, $\langle Af, f\rangle = \langle f, Af\rangle = \langle f, \lambda f\rangle = \bar{\lambda}\|f\|^2$, de modo que $\lambda = \bar{\lambda}$; es decir, λ es real. ■

En segundo lugar, *si f y g son autofunciones con autovalores λ y μ y $\lambda \neq \mu$, entonces f y g son ortogonales.*

Demostración Considérese $\langle Af, g\rangle - \langle f, Ag\rangle = 0 = \langle \lambda f, g\rangle - \langle f, \mu g\rangle = (\lambda, -\mu)\langle f, g\rangle$. Como $\mu \neq \lambda$, $\langle f, g\rangle = 0$. ■

Si f y g son autofunciones independientes con el mismo autovalor λ, entonces podemos obtener otras dos autofunciones ortogonales mediante el procedimiento de Gram-Schmidt. Se puede efectuar una operación similar con cualquier número de autofunciones. Así, *las autofunciones de un operador simétrico A forman un conjunto ortonormal.* He aquí la relación con las series de Fourier, pues si $\varphi_0, \varphi_1, \varphi_2, \ldots$ son las autofunciones ortonormales de A y, *si son completas,* entonces podemos escribir $\psi = \sum_{n=0}^{\infty}\langle\psi, \varphi_n\rangle\,\varphi_n$ para cualquier $\psi \in V$.

Por desgracia, es frecuente que las autofunciones no sean completas. Por ejemplo, el operador laplaciano en $\mathbb{R}^n$ no tiene autofunciones (de cuadrado integrable). Sin embargo, en muchos problemas importantes (problemas de "estados acotados"), las autofunciones son completas. Esto se demuestra en cursos avanzados de análisis funcional por medio de un teorema llamado el *teorema espectral.*

Regresemos a la interpretación física de los observables. De nuevo, sea V como antes y A un operador simétrico.

Intepretación física *Si A se "mide" sobre un estado ψ, sólo se observan los autovalores de A. El valor λ_n se observa con probabilidad $|\langle\psi, \varphi_n\rangle|^2$, donde $A\varphi_n = \lambda_n\varphi_n$.*

Esta interpretación es consistente, pues (véase el ejercicio 22 al final del capítulo)

$$1 = \langle\psi, \psi\rangle = \sum \langle\psi, \varphi_n\rangle\langle\varphi_n, \psi\rangle = \sum |\langle\psi, \varphi_n\rangle|^2,$$

de modo que la probabilidad total es 1. Además, si el sistema ya está en el estado φ_n (lo que, en general, no tiene que ocurrir), entonces observamos λ_n con probabilidad 1 (es decir, con total certidumbre). Así, el desarrollo de Fourier de ψ permite interpretar ψ como una "mezcla" de los autoestados φ_n, y los cuadrados de los valores absolutos de los coeficientes de Fourier son las probabilidades de observar los autovalores particulares.

Como ya hemos visto, en un estado dado ψ y dado un observable A, por lo general no se puede predecir con certidumbre el valor observado de A. Pero el promedio de los valores observados, después de muchos ensayos, es $\sum_{n=0}^{\infty} |\langle\psi, \varphi_n\rangle|^2\lambda_n = \langle A\psi, \psi\rangle$ (véase el ejercicio 3 de esta sección). Esta cantidad $\langle A\psi, \psi\rangle$ se llama ***valor esperado*** de A en el estado ψ. ¿Cuál es el valor esperado de un autoestado?

Probablemente el ejemplo más importante de un observable es el ***operador energía***, denotado H, llamado también ***hamiltoniano***. Para una sola partícula en un potencial U, está dado por

$$H\psi = -\frac{\hbar^2}{2m}\nabla^2\psi + U\psi.$$

La justificación de esta elección depende de un análisis más detallado de los fundamentos del tema; para ello referimos al lector a los textos de mecánica cuántica. En este caso, U es una función dada de valores que representa la energía potencial, m es la masa de la partícula y $\hbar$ es una cierta constante que depende de las unidades de medida $\hbar = 1.05 \times 10^{-27}$ erg · sec y que se denomina ***constante de Planck.***

Por ejemplo, en el átomo de hidrógeno sólo podemos observar niveles discretos de energía, que son los autovalores del operador

$$H\psi = -\frac{\hbar^2}{2m}\nabla^2\psi - \frac{\psi}{r},$$

donde $r(x, y, z) = (x^2 + y^2 + z^2)^{1/2}$ y m es la masa del electrón. Estos niveles de energía se miden directamente en los espectros atómicos. Hay que tener cuidado: cualquier ψ es un estado admisible, no solamente los autoestados; pero estos últimos son estados particularmente importantes, pues sus autovalores dan los valores que son observables.

La razón de la importancia del operador energía es doble. En primer lugar, sus autovalores dan las energías observables posibles. En segundo lugar, este operador determina la dependencia temporal de ψ por medio de la célebre ***ecuación de Schrödinger***:

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi$$

(la solución ψ de esta ecuación mediante series de Fourier aparece en el ejercicio 23 del final de este capítulo).

Otros operadores son:

1. El ***operador posición*** (en la dirección x):

$$Q_x(\psi) = x\psi(x, y, z).$$

2. El ***operador momento*** (en la dirección x):

$$P_x(\psi) = \frac{\hbar}{i}\,\frac{\partial\psi}{\partial x}.$$

3. El operador ***momento angular*** (en torno al eje z):

$$J_z(\psi) = \frac{\hbar}{i}\left(y\frac{\partial\psi}{\partial x} - x\frac{\partial\psi}{\partial y}\right).$$

Existen definiciones análogas para Q_y, Q_z, etcétera.

Las autofunciones de J_z son completas, y tanto éstas como sus correspondientes autovalores se calculan en cualquier libro de mecánica cuántica. Los operadores Q_x, P_x no tienen autofunciones de cuadrado integrable.

Por último, antes de pasar a un ejemplo específico, examinaremos el importante concepto de conmutador. El ***conmutador*** de dos operadores A y B es el operador $[A, B]$ definido por

$$[A, B] = AB - BA,$$

donde AB significa $A \circ B$, es decir, $(AB)(f) = A(B(f))$.

Supóngase que f es una autofunción de A y de B. Si $Af = \lambda f$ y $Bf = \mu f$, entonces

$$[A, B]\, f = (AB - BA)f = A(\mu f) - B(\lambda f) = \mu\lambda f - \lambda\mu f = 0.$$

Así, si A y B tienen las mismas autofunciones, entonces $[A, B]f = 0$ para dichas autofunciones. Si las autofunciones son completas, esperamos que $[A, B] = 0$ para todo f, pero esto requiere más hipótesis de las que analizaremos aquí.

Recíprocamente, si $[A, B] = 0$ *podemos elegir las autofunciones de modo que sean autofunciones simultáneas de A y B.* Esto es fácil de ver si no existen autovalores repetidos (el caso más general requiere de algunos argumentos adicionales). Como ejemplo, supóngase que $Af = \lambda f$. Entonces $A(Bf) - B(Af) = 0$, de modo que $A(Bf) - \lambda(Bf)$. Así, Bf es una autofunción de A, de modo que $Bf = \mu f$ para algún μ (ya que, por hipótesis, λ es un autovalor simple). Así, f es una autofunción de A y de B.

En resumen, $[A, B] = 0$ *sii A y B tienen autofunciones simultáneas.* Físicamente, esto significa que estas autofunciones dan observables exactos de A y B a la vez, o bien, diremos que A y B *se pueden medir simultáneamente.* Una justificación adicional de esta afirmación se da en el famoso ***principio de incertidumbre*** (ejercicio 5 de esta sección), que establece que el producto de los "errores" en la medida de A y B para el mismo ψ (es decir, medir A y B simultáneamente) es, al menos, $2|\langle C\psi, \psi\rangle|$, donde $C = [A, B]$. La definición de "error" también aparece en el ejercicio 5.

Por último, analicemos el ejemplo de una partícula en un "pozo infinitamente profundo". Otros ejemplos importantes, como el oscilador armónico ($H = -(\hbar^2/2m)\nabla^2 + kr$) y el átomo de hidrógeno ($H = -(\hbar^2/2m)\nabla^2 - 1/r$) se encuentran en los textos usuales, y son un poco más laboriosos. El hamiltoniano para el problema del pozo profundo en una dimensión es

$$H = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V\psi,$$

donde $V = 0$ en $[0, l]$ y $V = \infty$ fuera de dicho intervalo. Como esto no sirve para los cálculos dentro del espacio de las funciones de cuadrado integrable, vamos a reformular H pidiendo que

$$\begin{cases} H\psi = -\dfrac{\hbar^2}{2m}\dfrac{\partial^2\psi}{\partial x^2} & \text{en } [0, l] \\ \psi \ \text{ y } \ H\psi = 0 & \text{fuera de } [0, l]. \end{cases}$$

Así, H es realmente un operador sobre las funciones $\psi(x)$, $0 \leq x \leq l$, con $\psi(0) = \psi(l) = 0$.

Se sigue que ψ es una autofunción sii existe una constante E tal que

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} = E\psi.$$

La solución general de esta ecuación es

$$\psi(x) = A \operatorname{sen} \lambda x + B \cos \lambda x,$$

donde $\lambda^2 = 2mE/\hbar^2$. Si $\psi = 0$ en 0, l, entonces $B = 0$ también $\lambda = n\pi/l$, $n = 1, 2, \ldots$. Así, las autofunciones normalizadas son

$$\varphi_n(x) = \sqrt{2/l} \operatorname{sen} \lambda_n x$$

(véase la figura 10.9-2), donde $\lambda_n = n\pi/l$, $n = 1, 2, \ldots$, y los autovalores son

$$E_n = \frac{\hbar^2 n^2 \pi^2}{2ml^2}.$$

Así, estos E_n son los únicos valores de la energía que se pueden observar.

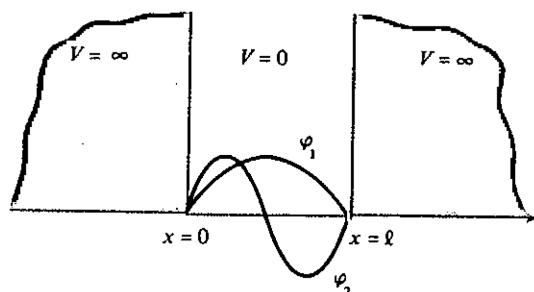


FIGURA 10.9-2 Los niveles de energía para una partícula en un pozo profundo

En este caso, estas funciones son completas, como demostramos en §10.3. Así, si una partícula está en un estado ψ, la probabilidad de que observemos la energía E_n está dada por

$$|\langle\psi,\varphi_n\rangle|^2 = \frac{2}{l}\left|\int_0^l \psi(x)\,\text{sen}\,\frac{n\pi x}{l}\,dx\right|^2$$

Ejercicios de §10.9

1. Sea A un operador en V. Defínase su ***adjunto*** A^* como $\langle A^*x, y\rangle = \langle x, Ay\rangle$ (supóngase que tal A^* existe). Demuéstrese que $(AB)^* = B^*A^*$.

2. Sea V igual a $\mathbb{R}^n$ y $A : V \to V$ lineal. Demuéstrese que A es simétrico en nuestro sentido si su matriz con respecto a cualquier base ortonormal es simétrica en el sentido usual de las matrices; es decir, si $a_{ij} + a_{ji}$.

3. Sea A un operador simétrico en V con un conjunto completo de autofunciones φ_0, φ_1, $\varphi_2, \ldots$ y autovalores $\lambda_0, \lambda_1, \ldots$. Si $\psi \in V$, muéstrese que el valor esperado de A está dado por

$$\langle A\psi,\psi\rangle = \sum_{n=0}^{\infty} |\langle\psi,\varphi_n\rangle|^2\lambda_n.$$

Interprétese esto en términos de la mecánica cuántica (es decir, desde el punto de vista de la probabilidad).

4. Supóngase que $\langle Af, g\rangle = -\langle f, Ag\rangle$ para todo $f, g \in V$. Muéstrese que A^2 es simétrico y que solamente tiene autovalores negativos.

5. ***Principio de incertidumbre.*** Sea A un operador simétrico. La ***incertidumbre*** (o ***varianza***) en la observación de A en un estado ψ está dada por

$$\Delta^2(A,\psi) = \langle (A - \langle A\psi,\psi\rangle)^2\psi,\psi\rangle.$$

a. Muéstrese que esto es igual a $\langle A^2\psi,\psi\rangle - \langle A\psi,\psi\rangle^2$.

b. Sean A, B dos operadores simétricos y $C = [A, B]$. Muéstrese que

$$\Delta^2(A,\psi)\Delta^2(B,\psi) \geq 4|\langle C\psi,\psi\rangle|^2.$$

Obsérvese el caso especial $[A, B] = 0$ e interprételo.

c. Para el caso $A\psi = x\psi$ y $B\psi = (\hbar^2/i)\partial\psi/\partial x$ (posición y momento), muéstrese que

$$\Delta^2(A, \psi)\Delta^2(B, \psi) \geq 4\hbar^2$$

(para $\|\psi\| = 1$). Éste es el ***principio de incertidumbre de Heisenberg.***

Demostraciones de los teoremas del capítulo 10

10.1.2 Teorema *El espacio V de las funciones continuas $f : [a, b] \to \mathbb{C}$ es un espacio con producto interno si definimos*

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\, dx.$$

Demostración Las propiedades del producto interno se siguen de estos cálculos:

i.

$$\begin{aligned}\langle af + bg, h\rangle &= \int_a^b [af(x) + bg(x)]\overline{h(x)}\, dx \\ &= a\int_a^b f(x)\overline{h(x)}\, dx + b\int_a^b g(x)\overline{h(x)}\, dx \\ &= a\langle f, g\rangle + b\langle f, g\rangle.\end{aligned}$$

ii.

$$\begin{aligned}\langle f, g\rangle &= \int_a^b f(x)\overline{g(x)}\, dx = \int_a^b \overline{\overline{f(x)}g(x)}\, dx \\ &= \overline{\int_a^b \overline{f(x)}g(x)\, dx} = \overline{\langle g, f\rangle}.\end{aligned}$$

iii. Obsérvese de **ii** que $\langle f, f\rangle = \overline{\langle f, f\rangle}$, de modo que $\langle f, f\rangle$ es real; así, $\langle f, f\rangle \geq 0$ tiene sentido. Entonces

$$\langle f, f\rangle = \int_a^b f(x)\overline{f(x)}\, dx = \int_a^b |f(x)|^2\, dx \geq 0,$$

pues $|f(x)|^2 \geq 0$.

Por último, supóngase que $\langle f, f\rangle = 0$. Si h es continua y $h \geq 0$, entonces

$$\int_a^b h(x)\,dx = 0 \quad \text{implica} \quad h = 0$$

(véase §8.4) y entonces $\langle f, f\rangle = 0$ implica $\int_a^b |f|^2 dx = 0$ y, por tanto, $f = 0$. ■

10.1.3 Desigualdad de Cauchy-Schwarz *Sean f, g pertenecientes al espacio con producto interno V.; entonces*

$$|\langle f, g\rangle| \leq \|f\|\,\|g\|.$$

Además, V con la norma $\|f\|$ satisface los axiomas de un espacio normado y, con $d(f, g) = \|f - g\|$, los axiomas de un espacio métrico.

Demostración En primer lugar, demostremos la desigualdad cuando $\|g\| = 1$:

$$\begin{aligned} 0 \leq \|f - \langle f, g\rangle g\|^2 &= \langle f - \langle f, g\rangle g, f - \langle f, g\rangle g\rangle \\ &= \langle f, f\rangle - \langle f, g\rangle\overline{\langle f, g\rangle} - \langle f, g\rangle\overline{\langle f, g\rangle} + \langle f, g\rangle\overline{\langle f, g\rangle}\langle g, g\rangle \\ &= \langle f, f\rangle - \langle f, g\rangle\overline{\langle f, g\rangle} = \|f\|^2 - |\langle f, g\rangle|^2. \end{aligned}$$

Así,

Para el caso general $|\langle f, g\rangle| \leq \|f\|\,\|g\|$, podemos suponer que $g \neq 0$, de modo que $\|g\| \neq 0$. Sea $h = g/\|g\|$, de modo que $\|h\| = 1$. Entonces $|\langle f, h\rangle| \leq \|f\|$. Pero $|\langle f, h\rangle| = |\langle f, g\rangle|/\|g\| \leq \|f\|$, por lo que obtenemos el resultado. ■

Este método es análogo al usado para demostrar la desigualdad de Cauchy-Schwarz en los espacios con producto interno real (capítulo 1), excepto que ahora es necesario un poco más de cuidado para llevar un registro de los complejos conjugados. El lector debe obtener las demás propiedades, teniendo un cuidado especial con la desigualdad triangular $\|f + g\| \leq \|f\| + \|g\|$.

10.16 Teorema *Sea $V = L^2$ el espacio de funciones $f: [a, b] \to \mathbb{C}$ tales que $|f|^2$ es integrable (es decir, $\int_a^b |f(x)|^2 < \infty$). Entonces el espacio V es un espacio con producto interno; con el producto interno dado por*

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\,dx$$

y la norma

$$\|f\| = \left(\int_a^b |f(x)|^2\,dx\right)^{1/2}$$

Demostración En primer lugar, si $\|f\| = 0$, tenemos que $\int_a^b |f(x)|^2 dx = 0$, de modo que por el teorema **8.3.4ii** del capítulo 8, f se anula excepto posiblemente en un conjunto de medida nula. Como estamos identificando las funciones que coinciden excepto en un conjunto de medida nula, $f = 0$. Por último, $\langle f, g\rangle$ satisface las demás reglas de un espacio con producto interno, como se muestra en el teorema **10.1.2**. Sólo necesitamos mostrar que $\langle f, g\rangle$ es finito (es decir, que fg es integrable).

Si sólo fuéramos a trabajar con funciones acotadas, fg sería integrable y acotada, pues lo son f y g (véase el capítulo 8). Como permitimos el uso de integrales impropias, f y g no tienen por qué estar acotadas. Si separamos f y g en sus partes real e imaginaria, y en sus partes positivas y negativas, podemos restringirnos al caso en que f y g sean reales y positivas (se pedirá al lector que desarrolle los detalles como ejercicio). Para cada $M \geq 1$, defínase $(fg)_M$ como en el capítulo 8. Queremos mostrar que

$$\lim_{M\to\infty} \int_a^b (fg)_M < \infty.$$

Al examinar los casos, vemos que

$$0 \leq (fg)_M \leq f_{\sqrt{M}} g_{\sqrt{M}} + f^2_{\sqrt{M}} + g^2_{\sqrt{M}}$$

de modo que

$$\begin{aligned}\int_a^b (fg)_M &\leq \langle f_{\sqrt{M}}, g_{\sqrt{M}}\rangle + \|f_{\sqrt{M}}\|^2 + \|g_{\sqrt{M}}\|^2 \\ &\leq \|f_{\sqrt{M}}\|^2 \|g_{\sqrt{M}}\|^2 + \|f_{\sqrt{M}}\|^2 + \|g_{\sqrt{M}}\|^2,\end{aligned}$$

por la desigualdad de Schwarz. Pero $\|f_{\sqrt{M}}\| \leq \|f\|$ y $\|g_{\sqrt{M}}\| \leq \|g\|$, de modo que

$$\int_a^b (fg)_M \leq \|f\|\,\|g\| + \|f\|^2 + \|g\|^2 < \infty.$$

En consecuencia obtenemos el resultado (los límites existen cuando la integral crece con M; sólo necesitábamos mostrar que estaba acotada superiormente).

Por último, para las funciones continuas a trozos, observemos que éstas forman un espacio vectorial (ejercicio 9 al final del capítulo) y que están acotadas (ejercicio 11). Por lo tanto, ambas funciones, f y $|f|^2$, son integrables, pues el conjunto de discontinuidades es finito (véase el capítulo 8). ■

10.2.1 Teorema *En un espacio con producto interno V, supóngase que $f = \sum_{k=0}^{\infty} c_k \varphi_k$ para una familia ortonormal $\varphi_0, \varphi_1, \ldots$ en V (convergencia en media) y que $f \in V$. Entonces $c_k = \langle f, \varphi_k\rangle = \overline{\langle \varphi_k, f\rangle}$.*

Demostración Sea $s_n = \sum_{k=0}^{n} c_k \varphi_k$, de modo que $\|f - s_n\| \to 0$. Fijemos i y sea $n \geq i$. Formamos

$$\langle f - s_n, \varphi_i\rangle = \langle f, \varphi_i\rangle - \langle s_n, \varphi_i\rangle.$$

Esta expresión tiende a cero cuando $n \to \infty$, pues $|\langle f - s_n, \varphi_i\rangle| \leq \|f - s_n\|$. Para $n \geq i$, tenemos

$$\langle s_n, \varphi_i\rangle = \sum_{k=0}^{n} \langle c_k\varphi_k, \varphi_i\rangle = \sum_{k=0}^{n} c_k\langle \varphi_k, \varphi_i\rangle = \sum_{k=0}^{n} c_k\delta_{ki} = c_i.$$

Así, $\langle f, \varphi_i\rangle - c_i \to 0$ cuando $n \to \infty$. Puesto que esta expresión es *independiente de* n, tenemos que $\langle f, \varphi_i\rangle = c_i$. ■

10.2.3 Desigualdad de Bessel *Sea* $\varphi_0, \varphi_1, \ldots$ *un sistema ortonormal en un espacio con producto interno V. Para cada* $f \in V$, *la serie real* $\sum_{i=0}^{\infty} |\langle f, \varphi_i\rangle|^2$ *converge, y tenemos la desigualdad*

$$\sum_{i=0}^{\infty} |\langle f, \varphi_i\rangle|^2 \leq \|f\|^2.$$

Demostración Sea $s_n = \sum_{i=0}^{n} \langle f, \varphi_i\rangle\varphi_i$. Primero mostraremos que $f - s_n$ y s_n son ortogonales. Para verlo, basta mostrar que $f - s_n$ y φ_i, $1 \leq i \leq n$, son ortogonales (¿por qué?). De hecho, $\langle f - s_n, \varphi_i\rangle = \langle f, \varphi_i\rangle - \langle s_n, \varphi_i\rangle$ y $\langle s_n, \varphi_i\rangle = \langle f, \varphi_i\rangle$, pues

$$\langle s_n, \varphi_i\rangle = \sum_{j=0}^{n} \langle\langle f, \varphi_j\rangle\varphi_j, \varphi_i\rangle = \sum_{j=0}^{n} \langle f, \varphi_j\rangle\delta_{ij} = \langle f, \varphi_i\rangle$$

(éste es el mismo cálculo que en la demostración del teorema **10.2.1**). Si g y h son ortogonales, entonces $\|g + h\|^2 = \|g\|^2 + \|h\|^2$ (teorema de Pitágoras, ejercicio 1, **§10.1**), y entonces

$$\|f\|^2 = \|f - s_n + s_n\|^2 = \|s_n\|^2 + \|f - s_n\|^2;$$

en consecuencia, $\|s_n\| \leq \|f\|^2$. Ahora

$$\|s_n\|^2 = \left\|\sum_{i=0}^{n} \langle f, \varphi_i\rangle\varphi_i\right\|^2 = \sum_{i=0}^{n} |\langle f, \varphi_i\rangle|^2\|\varphi_i\|^2,$$

pues las φ_i son ortogonales, y en consecuencia,

$$\|s_n\|^2 = \sum_{i=0}^{n} |\langle f, \varphi_i\rangle|^2 \leq \|f\|^2.$$

Así, la suma parcial de la serie $\sum_{i=0}^{\infty} |\langle f, \varphi_i\rangle|$ es $\|s_n\|^2$, que es una sucesión creciente, pues los términos de la serie son ≥ 0 y la serie está acotada superiormente por $\|f\|^2$. Por lo tanto, la serie converge a una suma $\leq \|f\|^2$. ■

10.2.4 Teorema de Parseval *Sea V un espacio con producto interno y φ_0, $\varphi_1, \ldots$ un sistema ortonormal. Entonces $\varphi_0, \varphi_1, \ldots$ es completo sii para cada $f \in V$ tenemos*

$$\|f\|^2 = \sum_{n=0}^{\infty} |\langle f, \varphi_n \rangle|^2.$$

Demostración Sea $s_n = \sum_{i=0}^{n} \langle f, \varphi_i \rangle \varphi_i$. En la demostración de la desigualdad de Bessel se mostró que

$$\|f\|^2 = \|f - s_n\|^2 + \|s_n\|^2.$$

Si $\varphi_0, \varphi_1, \ldots$ es completo, entonces $s_n \to f$, por lo que $\|f - s_n\|^2 \to 0$. En consecuencia, si $n \to \infty$ en $\|s_n\|^2 = \sum_{i=0}^{n} |\langle f, \varphi_i \rangle|^2$ obtenemos $\|f\|^2 = \sum_{i=0}^{\infty} |\langle f, \varphi_i \rangle|^2$.

Recíprocamente, si se cumple esta relación, entonces $\|f\|^2 - \|s_n\|^2 \to 0$ cuando $n \to \infty$, y entonces $\|f - s_n\|^2 \to 0$; es decir, $s_n \to f$, lo que significa que

$$f = \sum_{i=0}^{\infty} \langle f, \varphi_i \rangle \varphi_i. \quad \blacksquare$$

10.2.5 Teorema de la mejor aproximación media *Sea V un espacio con producto interno y $\varphi_0, \varphi_1, \ldots, \varphi_n$ un conjunto de vectores ortonormales en V. Entonces, para cada conjunto de números $t_0, t_1, \ldots, t_n$,*

$$\left\| f - \sum_{k=0}^{n} t_k \varphi_k \right\| \geq \left\| f - \sum_{k=0}^{n} \langle f, \varphi_k \rangle \varphi_k \right\|.$$

La igualdad se cumple sii $t_k = \langle f, \varphi_k \rangle$.

Demostración Sean $c_k = \langle f, \varphi_k \rangle$, $s_n = \sum_{i=0}^{n} c_i \varphi_i$, y $h_n = \sum_{i=0}^{n} t_i \varphi_i$. Se pide mostrar que

$$\|f - s_n\|^2 \leq \|f - h_n\|^2$$

y que la igualdad se da sii $t_k = c_k$. Para ver esto mostraremos que

$$\|f - h_n\|^2 = \|f\|^2 - \sum_{k=0}^{n} |c_k|^2 + \sum_{k=0}^{n} |c_k - t_k|^2 = \|f - s_n\|^2 + \sum_{k=0}^{n} |c_k - t_k|^2,$$

lo que basta para demostrar el teorema. Para demostrar esta igualdad, obsérvese que

$$\|f - h_n\|^2 = \langle f - h_n, f - h_n \rangle = \langle f, f \rangle - \langle f, h_n \rangle - \langle h_n, f \rangle + \langle h_n, h_n \rangle.$$

En primer lugar,

$$\langle h_n, h_n\rangle = \sum_{i,j}\langle t_i\varphi_i, t_j\varphi_j\rangle = \sum_{i,j} t_i\overline{t_j}\delta_{ij} = \sum_{i=0}^{n} |t_i|^2.$$

en segundo lugar,

$$\langle f, h_n\rangle = \left\langle f, \sum_{k=0}^{n} t_k\varphi_k \right\rangle = \sum_{k=0}^{n} c_k\overline{t_k}.$$

así,

$$\|f - h_n\|^2 = \|f^2\| - \sum_{k=0}^{n} c_k\overline{t_k} - \sum_{k=0}^{n} t_k\overline{c_k} + \sum_{k=0}^{n} |t_k|^2 = \|f\|^2 - \sum_{k=0}^{n} |c_k|^2 + \sum_{k=0}^{n} |c_k - t_k|^2,$$

como se pedía. ■

10.3.1 Teorema de completitud media *Los sistemas exponencial y trigonométrico en* $[0, 2\pi]$ *(o* $[-\pi, \pi]$*) son completos en el espacio* $V = L^2$ *de funciones* f : $[0, 2\pi] \to \mathbb{C}$ *con* $\int_0^{2\pi} |f(x)|^2 dx < \infty$ *(la integral puede ser impropia).*

Demostración[4] Por nuestras observaciones en el texto y en el ejercicio 1 de **§10.3**, basta considerar el caso exponencial. Los siguientes lemas contienen dos resultados necesarios.

Lema 1 *(Teorema de Stone-Weierstrass en un caso particular) Sea* $f : [0, 2\pi] \to \mathbb{C}$ *continua y* $f(0) = f(2\pi)$ *(periodicidad). Entonces, para cada* $\varepsilon > 0$ *existe* n *y existen constantes* c_i, $i = -n, \ldots, -1, 0, 1, \ldots, n$, *tales que si formamos la función*

$$p_n(x) = c_0 + c_1 e^{ix} + c_2 e^{2ix} + \cdots + c_n e^{nix} + c_{-1} e^{-ix} + c_{-2} e^{-2ix} + \cdots + c_{-n} e^{-nix},$$

entonces

$$|f(x) - p_n(x)| < \varepsilon$$

para todo $x \in [0, 2\pi]$.

[4] Una demostración de Luxemburg que no está basada en el teorema de Stone-Weierstrass se apunta en el ejercicio 75 del final del capítulo. Otra demostración, debida a Lebesgue, aparece en el ejercicio 76. Sin embargo, ambas demostraciones se basan en el recíproco del ejemplo **10.2.7** (véase el ejercicio 14 al final del capítulo), que usa la completitud de L^2; es decir, la integral de Lebesgue.

En el capítulo 5 demostramos el teorema de Stone-Weierstrass. Véase también el ejercicio 44b, capítulo 5.

Lema 2 *Sea $f: [0, 2\pi] \to \mathbb{C}$ de cuadrado integrable y $\varepsilon > 0$. Entonces existe una función continua $g: [0, 2\pi] \to \mathbb{C}$ tal que $g(0) = g(2\pi)$ y*

$$\|f - g\|^2 = \int_0^{2\pi} |f(x) - g(x)|^2\,dx < \varepsilon.$$

Demostración Supóngase primero que $f \geq 0$ y que está acotada superiormente por M. Dado $\varepsilon > 0$, elegimos una partición P de $[0, 2\pi]$ tal que, si $h = f^2$,

$$\left|\int_0^{2\pi} h - \sum_{i=1}^{n} h(c_i)(x_{i+1} - x_i)\right| < \frac{\varepsilon}{2},$$

y hacemos una estimación similar para f. Podemos trazar una gráfica compuesta por líneas rectas correspondiente a una función continua g que sea constante $= f(c_i)$ en $[y_i, z_i]$, donde $[y_i, z_i] \subset [x_i, x_{i+1}]$, $|y_i - x_i| < \varepsilon/8M^2n$ y $|x_{i+1} - z_i| < \varepsilon/8M^2n$, y tal que g esté acotada por M. Entonces

$$\int_0^{2\pi} |f - g|^2\,dx = \int_0^{2\pi} (f^2 + g^2 - 2fg)\,dx < \frac{\varepsilon}{2} + 4M^2 \cdot \frac{\varepsilon}{8M^2n} \cdot n = \varepsilon,$$

sumando y restando las aproximaciones de $\int f^2 = \int h$ y $\int f$ y usando la definición de g. Los detalles se dejan al lector.

Se puede tratar el caso general si escribimos f como $f = f^+ - f^-$ (véase el capítulo 8), de modo que podemos suponer que $f \geq 0$. Formamos f_M como en el capítulo 8 y elegimos M suficientemente grande como para que $\int |f - f_M|^2 < \varepsilon/4$, lo que es posible por un corolario del teorema de la convergencia monótona (capítulo 8). Así, podemos elegir g tal que $\int |g - f_M|^2 < \varepsilon/4$, y entonces $\int |f - g|^2 < \varepsilon$, pues

$$\|f - g\| \leq \|f - f_M\| + \|g - f_M\| \leq \frac{\sqrt{\varepsilon}}{2} + \frac{\sqrt{\varepsilon}}{2} = \sqrt{\varepsilon}. \quad \blacksquare$$

Para demostrar el teorema a partir de estos lemas procederemos en dos pasos.

Paso 1 *Demostración del teorema para f continua y periódica.*

Sea

$$s_n = \sum_{-n}^{n} \langle f, \varphi_k \rangle \varphi_k, \quad \text{donde } \varphi_k(x) = \frac{e^{ixk}}{\sqrt{2\pi}}.$$

Para $\varepsilon > 0$, debemos mostrar que existe N tal que $n \geq N$ implica $\|f - s_n\| < \varepsilon$. Basta producir un solo n, pues, por el teorema **10.2.5**, $\|f - s_{n+k}\| \leq \|f - s_n\|$. Usamos el lema **1** para elegir p_n tal que $|f(x) - p_n(x)| < \varepsilon/\sqrt{2\pi}$, y formamos el s_n correspondiente. Ahora,

$$\|f - p_n\|^2 = \int_0^{2\pi} |f(x) - p_n(x)|^2 \, dx \leq \int_0^{2\pi} \left(\frac{\varepsilon}{\sqrt{2\pi}}\right)^2 dx = \varepsilon^2.$$

Así, $\|f - p_n\| < \varepsilon$. Por el teorema **10.2.5**,

$$\|f - s_n\| \leq \|f - p_n\| < \varepsilon,$$

pues la serie de Fourier da la mejor aproximación media a f. Esto demuestra el paso **1.**

Paso 2 *Caso general.*

En vista del lema **2** y el paso **1**, basta demostrar lo siguiente. En este caso, V es el espacio de las funciones de cuadrado integrable, pero el lema se enuncia en términos generales.

Lema 3 *Sea V un espacio con producto interno y sea $\varphi_0, \varphi_1, \ldots$ una familia ortonormal. Supóngase que $f \in V$ y $f_n \to f$. Si*

$$f_n = \sum_{k=0}^{\infty} \langle f_n, \varphi_k \rangle \varphi_k$$

para cada n, entonces

$$f = \sum_{k=0}^{\infty} \langle f, \varphi_k \rangle \varphi_k.$$

Demostración Dado $\varepsilon > 0$, sea N tal que $k \geq N$ implica $\|f_k - f\| = \varepsilon/3$. Elijamos M tal que $n \geq M$ implique

$$\left\| \sum_{j=0}^{n} \langle f_N, \varphi_j \rangle \varphi_j - f_N \right\| < \frac{\varepsilon}{3}.$$

Usando la desigualdad triangular,

$$\left\|\sum_{j=0}^{n}\langle f,\varphi_j\rangle\varphi_j - f\right\| \le \left\|\sum_{j=0}^{n}\langle f,\varphi_j\rangle\varphi_j - \sum_{j=0}^{n}\langle f_N,\varphi_j\rangle\varphi_j\right\| + \left\|\sum_{j=0}^{n}\langle f_N,\varphi_j\rangle\varphi_j - f_N\right\| + \|f_N - f\|.$$

Por la desigualdad de Bessel, el primer término es $\le \|f - f_N\|$. Así, $n \ge M$ implica

$$\left\|\sum_{j=0}^{n}\langle f,\varphi_j\rangle\varphi_j - f\right\| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

lo que demuestra nuestra afirmación. ■

10.3.2 Teorema de la convergencia puntual *Sea $f: [0, 2\pi] \to \mathbb{R}$ (o $f: [-\pi, \pi] \to \mathbb{R}$) continua a trozos, con una discontinuidad de salto en x_0, y supóngase que $f'(x_0^+)$ y $f'(x_0^-)$ existen. Entonces la serie de Fourier de f (en forma exponencial o trigonométrica) evaluada en x_0 converge a $[f(x_0^+) + f(x_0^-)]/2$. En particular, si f es derivable en x_0, la serie de Fourier de f converge en x_0 a $f(x_0)$.*

Demostración Es conveniente demostrar primero el siguiente caso particular:

Lema 4 *Sea $f:[-\pi,\pi] \to \mathbb{C}$ de cuadrado integrable y derivable en x_0 (como siempre, extendemos f de modo que sea periódica). Entonces la serie de Fourier de f en x_0 converge a $f(x_0)$.*

Demostración (Seguimos una demostración de P. Chernoff.) Si trasladamos y sumamos una constante, podemos suponer que $x_0 = 0$ y que $f(x_0) = 0$ (¿por qué?). Defínase una nueva función $g(x)$ como

$$g(x) = \frac{f(x)}{e^{ix} - 1} \text{ si } x \neq 0 \quad \text{y} \quad g(x) = \frac{f'(0)}{i} \text{ si } x = 0.$$

Por la regla del cociente del cálculo se sigue que g es continua en 0. Como $1/(e^{ix} - 1)$ está acotada en valor absoluto fuera de una vecindad de 0, se sigue que g es de cuadrado integrable (¿por qué?).

Ahora, $f(x) = (e^{ix} - 1)\, g(x)$. Sea $c_n(f)$ el n-ésimo coeficiente de Fourier de f y $c_n(g)$ el coeficiente correspondiente de g. Por definición,

$$c_n(f) = c_{n-1}(g) - c_n(g).$$

Así,

$$\sum_{n=-N}^{N} c_n(f) = c_{-N-1}(g) - c_N(g),$$

pues tenemos una suma telescópica. Como $x_0 = 0$, $\sum_{-N}^{N} c_n(f)$ es la N-ésima suma parcial en $x = 0$ de la serie de Fourier de f. Pero $c_N(g) \to 0$, por la desigualdad de Bessel. En consecuencia, $\sum_{-N}^{N} c_n(f) \to 0 = f(x_0)$. ■

En realidad, no necesitamos el hecho de que f sea derivable en x_0. Si f "es Lipschitz" en x_0 (es decir, si existe una constante M tal que $|f(x) - f(x_0))/(x - x_0)| \leq M$ para $|x - x_0| < \delta$, $x \neq x_0$), podríamos obtener el mismo resultado mediante una demostración análoga (sólo necesitamos que la función g de la demostración sea de cuadrado integrable, o incluso solamente integrable). Por ejemplo, si f es continua y $f'(x_0^+)$ y $f'(x_0^-)$ existen, entonces se satisface esta condición (¿por qué?).

Para demostrar el teorema a partir del lema **4** y las observaciones anteriores, considérese

$$h(x) = \begin{cases} f(x_0^-) & x < x_0 \\ f(x_0) & x = x_0 \\ f(x_0^+) & x > x_0. \end{cases}$$

Entonces h es una función escalonada y podemos calcular su serie de Fourier de forma explícita (como en el apartado 1, tabla 10.5-4; véase el ejemplo **10.5.3**). En particular, esta serie converge a $[f(x_0^-) + f(x_0^+)]/2$ en x_0. Considérese ahora la función

$$k(x) = f(x) - h(x).$$

Entonces $k(x_0) = 0 = k(x_0^+) = k(x_0^-)$ y $k'(x_0^+)$, $k'(x_0^-)$ existen. De donde, por el lema **4**, la serie de Fourier de k converge a 0 en x_0. En consecuencia, la serie de Fourier de f converge a $[f(x_0^+) + f(x_0^-)]/2$ en x_0. Esto demuestra la afirmación. ■

Pasemos ahora a la larga demostración clásica del teorema **10.3.2**. Más adelante será conveniente tener esta demostración a mano, a pesar de que es más complicada que la recién expuesta. En primer lugar, explicaremos la idea básica detrás de la demostración. Sea $s_n(x)$ la n-ésima suma parcial de la serie de Fourier trigonométrica. Escribiremos

$$s_n(x) = \int_0^{2\pi} f(\xi) D_n(x - \xi)\, d\xi$$

para una función D_n que especificaremos más adelante (lema **9**); decimos que s_n es la ***convolución*** de f y D_n. Demostraremos que D_n tiene área unidad y que se "concentra" en torno a 0; es decir, se comporta como una función delta de Dirac. Cuando $n \to \infty$, la convolución elegirá el valor de f en x. Véase la figura 10.P-1. Por esta razón, también diremos que D_n es una ***identidad aproximada***.

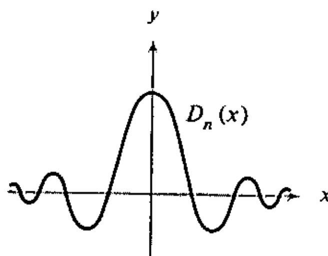


FIGURA 10.P-1 La gráfica de la función identidad aproximada $D_n(x)$

Antes de poder formalizar estas ideas, necesitamos algunos resultados preliminares. El primer lema es una generalización del ejemplo **10.2.9.**

Lema 5 Lema de Riemann-Lebesgue *Supóngase que f está acotada y es integrable (Riemann) en* $[a, b]$. *Entonces*

$$\lim_{\alpha \to \infty} \int_a^b f(x) \operatorname{sen}(\alpha x)\, dx = 0$$

(donde el límite se toma para los reales $\alpha > 0$).

Demostración Primero supóngase que f es una constante M. Entonces

$$\left| \int_a^b f(x) \operatorname{sen} \alpha x\, dx \right| = |M| \left| \int_a^b \operatorname{sen} \alpha x\, dx \right| = |M| \frac{|\cos(\alpha a) - \cos(\alpha b)|}{\alpha}$$

$$\leq \left| \frac{2M}{\alpha} \right| \to 0 \text{ cuando } \alpha \to \infty.$$

Así, el resultado es cierto si f es una constante.

Para el caso general, dado $\varepsilon > 0$ elegimos una partición $P = \{x_0, x_1, \ldots, x_n\}$ de $[a, b]$ tal que $U(f, P) - L(f, P) < \varepsilon/2$. Entonces

$$U(f, P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1}) \quad \text{y} \quad L(f, P) = \sum_{i=1}^{n} m_i(x_i - x_{i-1}),$$

donde M_i es el supremo de f en $[x_{i-1}, x_i]$ y m_i es el ínfimo. Sea m la función escalonada igual a m_i en $]x_{i-1}, x_i[$. Elegimos N tal que

$$\int_a^b m(x)\operatorname{sen}\alpha x\,dx = \sum_{i=1}^{n}\int_{x_{i-1}}^{x_i} m_i \operatorname{sen}\alpha x\,dx < \frac{\varepsilon}{2}$$

si $\alpha \geq N$, lo que es posible pues m_i es constante y n es fijo y finito. Por la desigualdad triangular, para $\alpha \geq N$,

$$\left|\int_a^b f(x)\operatorname{sen}\alpha x\,dx\right| \leq \left|\int_a^b m(x)\operatorname{sen}\alpha x\,dx\right| + \left|\int_a^b [f(x) - m(x)]\operatorname{sen}\alpha x\,dx\right|$$
$$< \frac{\varepsilon}{2} + \int_a^b |M(x) - m(x)|\,dx,$$

donde $M = M_i$ en $]x_{i-1}, x_i]$ (aquí hemos utilizado $|\operatorname{sen}\alpha x| \leq 1$). Pero $M(x) - m(x) \geq 0$ y

$$\int_a^b [M(x) - m(x)]\,dx = U(f,P) - L(f,P) < \frac{\varepsilon}{2},$$

y entonces. para $\alpha > N$,

$$\left|\int_a^b f(x)\operatorname{sen}\alpha x\,dx\right| < \varepsilon. \quad \blacksquare$$

Lema 6 *Supóngase que $g : [0, a] \to \mathbb{R}$ es continua a trozos y que $g'(0^+)$ existe. Entonces*

$$\lim_{k\to\infty}\int_0^a g(x)\frac{\operatorname{sen}kxdx}{x} = \frac{\pi}{2}g(0^+).$$

Demostración. Como

$$\int_0^a g(x)\frac{\operatorname{sen}kxdx}{x} = g(0^+)\int_0^a \frac{\operatorname{sen}kxdx}{x} + \int_0^a \frac{g(x) - g(0^+)}{x}\operatorname{sen}kx\,dx,$$

basta mostrar que

$$\int_0^a \frac{\operatorname{sen}kxdx}{x}dx \to \frac{\pi}{2} \quad \text{cuando } k \to \infty, \tag{1}$$

y que

$$\int_0^a \frac{g(x) - g(0^+)}{x}\operatorname{sen}kx\,dx \to 0 \quad \text{cuando } k \to \infty. \tag{2}$$

Para demostrar (1), obsérvese que

$$\int_0^a \frac{\operatorname{sen} kx}{x}\, dx = \int_0^{ka} \frac{\operatorname{sen} t}{t}\, dt,$$

lo que converge a $\pi/2$ cuando $k \to \infty$, pues $\int_0^\infty [(\operatorname{sen} t)/t]dt = \pi/2$; véanse el ejemplo **8.5.6** y el ejercicio 29 al final de este capítulo.

Para demostrar la ecuación (2), obsérvese que $[g(x) - g(0^+)]/x$ está acotada y es integrable (pues cuando $x \to 0$ esto tiende a un límite $g'(0^+)$). En consecuencia,

$$\int_0^a \frac{g(x) - g(0^+)}{x} \operatorname{sen} kx\, dx \to 0$$

cuando $k \to \infty$, por el lema **5.** ■

Necesitamos el lema **5** para α real y $[a, b]$ un intervalo arbitrario. Obsérvese que este caso no se sigue del ejemplo **10.2.9.**

Lema 7 *Sea g continua a trozos en $]a, b[$ y con una discontinuidad de salto en x_0. Supóngase que $g'(x_0^+)$ y $g'(x_0^-)$ existen. Entonces*

$$\lim_{k\to\infty} \int_a^b g(x)\frac{\operatorname{sen} k(x - x_0)}{x - x_0}\, dx = \frac{\pi[g(x_0^+) + g(x_0^-)]}{2}.$$

Demostración Escribimos la integral como una suma,

$$\int_a^b = \int_a^{x_0} + \int_{x_0}^b,$$

y observamos que

$$\int_a^{x_0} g(x)\frac{\operatorname{sen}[k(x - x_0)]}{x - x_0}\, dx = \int_0^{x_0 - a} g(x_0 - t)\frac{\operatorname{sen} kt}{t}\, dt$$

y que

$$\int_{x_0}^b g(x)\frac{\operatorname{sen}[k(x - x_0)]}{x - x_0}\, dx = \int_0^{b - x_0} g(x_0 - t)\frac{\operatorname{sen} kt}{t}\, dt.$$

Ahora, sea $k \to \infty$ y utilicemos el lema **6.** Obtenemos $\pi g(x_0^-)/2$ y $\pi g(x_0^+)/2$ como límite de estas dos integrales, respectivamente, cuando $k \to \infty$. ■

Lema 8 *Sea* $f: [0, 2\pi] \to \mathbb{R}$. *Entonces la n-ésima suma parcial de la serie de Fourier de f se puede escribir como*

$$s_n(x) = \frac{1}{2\pi} \int_0^{2\pi} f(t)\, dt + \frac{1}{\pi} \sum_{k=1}^{n} \int_0^{2\pi} f(t) \cos k(t-x)\, dt.$$

Demostración Esto está claro si recordamos que

$$\cos[k(t-x)] = \cos kt \cos kx + \operatorname{sen} kt \operatorname{sen} kx. \quad \blacksquare$$

Lema 9 *Sea* $s_n(x)$ *la n-ésima suma parcial de la serie de Fourier de f. Entonces*

$$s_n(x) = \frac{1}{2\pi} \int_0^{2\pi} f(t) D_n(t-x)\, dt \quad \text{donde} \quad D_n(u) = \frac{\operatorname{sen}[(n+1/2)u]}{\operatorname{sen}(u/2)}.$$

Demostración Esto es consecuencia del lema **8** y de la identidad

$$\sum_{k=-n}^{n} e^{iku} = 1 + 2\sum_{k=1}^{n} \cos ku = \frac{\operatorname{sen}[(n+1/2)u]}{\operatorname{sen}(u/2)}$$

(ejercicio 5, §**10.2**). $\blacksquare$

Ahora estamos listos para demostrar el teorema de convergencia puntual. Debemos mostrar que

$$s_n(x_0) \to \frac{f(x_0^+) + f(x_0^-)}{2}$$

cuando $n \to \infty$. Supondremos que $0 < x_0 < 2\pi$. Pedimos al lector que analice los casos $x_0 = 0$ y $x_0 = 2\pi$ por separado. Por el lema **9**,

$$s_n(x_0) = \frac{1}{\pi} \int_0^{2\pi} g(t) \frac{\operatorname{sen}[(n+1/2)(t-x_0)]}{t-x_0}\, dt,$$

donde

$$g(x) = f(t) \frac{(t-x_0)/2}{\operatorname{sen}[(t-x_0)/2]}, \quad x_0 \neq t.$$

Por el lema **7** (que es aplicable, por el ejercicio 41 al final de este capítulo), tenemos

$$s_n(x_0) \to \frac{g(x_0^+) + g(x_0^-)}{2}.$$

Ahora es muy fácil ver que

$$g(x_0^+) = f(x_0^+) \quad \text{y} \quad g(x_0^-) = f(x_0^-),$$

con lo que obtenemos el teorema. ■

10.4.2 Teorema de Fejér *Supóngase que f es continua a trozos en $[0, 2\pi]$ y que $f(x_0^+)$ y $f(x_0^-)$ existen. Entonces la serie de Fourier de f converge $(C, 1)$ en x_0 a $[f(x_0^+) + f(x_0^-)]/2$. Si f es continua y $f(0) = f(2\pi)$, la serie de Fourier converge $(C, 1)$ uniformemente a f.*

Demostración Por razones de notación, es un poco más conveniente usar $[-\pi, \pi]$ en vez de $[0, 2\pi]$ para esta demostración; esto no afecta las conclusiones. Considérese $s_n(x) = \sum_{k=-n}^{n} c_k e^{ikx}$, la n-ésima suma parcial de la serie de Fourier de f. Para analizar la sumabilidad $(C, 1)$, sea

$$\sigma_n(x) = \frac{1}{n+1} \sum_{k=0}^{n} s_k(x).$$

Por el lema **9**,

$$\sigma_n(x) = \frac{1}{n+1} \sum_{k=0}^{n} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) D_k(t)\, dt;$$

es decir,

$$\sigma_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) F_n(t)\, dt,$$

donde

$$F_n(t) = \frac{1}{n+1} \sum_{k=0}^{n} D_k(t)$$

es el ***núcleo de Fejér***. Necesitaremos los siguientes lemas.

Lema 10 $F_n(t) = \dfrac{1}{n+1} \dfrac{\operatorname{sen}^2[(n+1)t/2]}{\operatorname{sen}^2[t/2]}.$

Demostración Por la fórmula para D_n (lema **9**), tenemos

$$(n+1)F_n(t) = \sum_{k=0}^{n} \frac{\operatorname{sen}[(k+1/2)t]}{\operatorname{sen}(t/2)} = \frac{1}{\operatorname{sen}(t/2)} \operatorname{Im}\left\{\sum_{k=0}^{n} e^{i(k+1/2)t}\right\}$$

$$= \frac{1}{\operatorname{sen}(t/2)} \operatorname{Im}\left\{e^{it/2}\frac{e^{i(n+1)t}-1}{e^{it}-1}\right\}$$

$$= \frac{1}{\operatorname{sen}(t/2)} \operatorname{Im}\left\{\frac{e^{i(n+1)t}-1}{e^{it/2}-e^{-it/2}}\right\}$$

$$= \frac{1-\cos[(n+1)t]}{2\operatorname{sen}^2(t/2)} = \frac{\operatorname{sen}^2[1/2(n+1)t]}{\operatorname{sen}^2(t/2)}. \quad \blacksquare$$

Lema 11 *El núcleo de Fejér tiene las siguientes propiedades:*

i. $F_n(t)$ *es* 2π*-periódico.*

ii. $\dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} F_n(t)\,dt = 1.$

iii. $F_n(t) \geq 0.$

iv. *Para cada* $\delta > 0$ *fijo,* $\lim_{n\to\infty} \int_{\delta\leq|t|\leq\pi} F_n(t)\,dt = 0.$

Demostración **i** y **ii** se siguen de la definición de F_n; **iii** es consecuencia del lema **10**. **iv:** Para $\delta \leq |t| \leq \pi$ tenemos $1/(\operatorname{sen}^2 t/2) \leq (\operatorname{sen}^2 \delta/2)$. Por lo tanto,

$$0 \leq F_n(t) \leq \frac{1}{n+1}\frac{1}{\operatorname{sen}^2(\delta/2)}, \quad \delta \leq |t| \leq \pi.$$

Como esto converge a 0 uniformemente cuando $n \to \infty$, la integral $\int_{\delta\leq|t|\leq\pi} F_n(t)dt \to 0.$ $\blacksquare$

Vamos a demostrar ahora el teorema de Fejér. Usamos la misma técnica que en la demostración del teorema de la convergencia puntual (véanse los argumentos posteriores al lema **4**), por lo que basta demostrar la última parte del teorema. Así, supóngase que f es continua. Tenemos

$$\sigma_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)F_n(t)\,dt.$$

Por lo tanto, por **ii** del lema **11**,

$$f(x) - \sigma_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} (f(x) - f(x-t))F_n(t)\,dt.$$

De acuerdo con esto, por **iii** del lema **11** (positividad de F_n),

$$|f(x) - \sigma_n(x)| \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - f(x-t)|F_n(t)\,dt.$$

Dado $\varepsilon > 0$, podemos usar la continuidad uniforme de f para determinar $\delta > 0$ tal que $|f(x) - f(y)| \le \varepsilon$ si $|x - y| \le \delta$. Entonces

$$\begin{aligned}|f(x) - \sigma_n(x)| \le\ & \frac{1}{2\pi}\int_{|t|\le\delta} |f(x) - f(x-t)|F_n(t)\,dt \\ & + \frac{1}{2\pi}\int_{\delta\le|t|\le\pi} |f(x) - f(x-t)|F_n(t)\,dt.\end{aligned}$$

La primera integral del miembro derecho está acotada superiormente por

$$\frac{1}{2\pi}\int_{|t|\le\delta} \varepsilon F_n(t)\,dt \le \frac{1}{2\pi}\int_{-\pi}^{\pi} \varepsilon F_n(t)\,dt = \varepsilon.$$

La segunda integral está acotada por

$$\frac{1}{2\pi}\int_{\delta\le|t|\le\pi} 2MF_n(t)\,dt = \frac{M}{\pi}\int_{\delta\le|t|\le\pi} F_n(t)\,dt,$$

donde $M = \sup_t |f(t)|$. Por la propiedad **iv** del lema **11**, podemos elegir N de modo que si $n \ge N$, esta última integral sea $\le \varepsilon$. Así, si $n \ge N$, $|f(x) - \sigma_n(x)| \le \varepsilon + \varepsilon = 2\varepsilon$. ■

Se puede demostrar que las sumas de Cesaro convergen a una función integrable, excepto posiblemente en un conjunto de medida nula (véase Hewitt y Stromberg, *Real and Abstract Analysis*, Springer-Verlag, pág. 294).

10.5.1 Teorema de integración *Supóngase que $\int_{-\pi}^{\pi} |f(x)|^2\,dx < \infty$ y que f tiene la serie de Fourier*

$$\frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \operatorname{sen} nx).$$

Entonces, si $g(x) = \int_{-\pi}^{x} f(y)\,dy$, tenemos

$$\begin{aligned} g(x) &= \frac{a_0(x+\pi)}{2} + \sum_{n=1}^{\infty}\left(a_n\int_{-\pi}^{x}\cos ny\,dy + b_n\int_{-\pi}^{x}\operatorname{sen} ny\,dy\right) \\ &= \frac{a_0(x+\pi)}{2} + \sum_{n=1}^{\infty}\left\{\frac{a_n}{n}\operatorname{sen} nx + \frac{b_n}{n}((-1)^n - \cos nx)\right\}\end{aligned}$$

y la convergencia es uniforme para $-\pi \le x \le \pi$.

Demostración Prepararemos el siguiente lema.

Lema 12 *Supóngase que $f_n : [a, b] \to \mathbb{R}$ son tales que $\int_a^b |f_n(x)|^2 \, dx < \infty$ y $f_n \to f$ en media. Sean*

$$g_n(x) = \int_a^x f_n(y)\, dy \qquad \text{y} \qquad g(x) = \int_a^x f(y)\, dy.$$

Entonces $g_n \to g$ uniformemente en $[a, b]$.

Demostración Por la desigualdad de Cauchy-Schwarz,

$$\begin{aligned} |g_n(x) - g(x)|^2 &\leq \left(\int_a^x |f_n(y) - f(y)|\, dy \right)^2 \\ &\leq \left(\int_a^x |f_n(y) - f(y)|^2\, dy \right) (x - a) \leq \|f_n - f\|^2 (b - a), \end{aligned}$$

de donde es obvio el resultado. ■

Para el teorema, sea $s_n(x)$ la n-ésima suma parcial de la serie de Fourier y sea $f_n = s_n$ en el lema. Sabemos que $f_n \to f$ en media (teorema **10.3.1**) y que $g_n \to g$ uniformemente. En este caso, g_n es la suma parcial de la serie de Fourier integrada, por lo que obtenemos el resultado. ■

10.5.2 Fenómeno de Gibbs *Considérese*

$$f(x) = \begin{cases} a & -\pi \leq x < 0 \\ b & 0 \leq x \leq \pi, \end{cases}$$

y supóngase que $a < b$. Sea $s_n(x)$ la n-ésima suma parcial de la serie de Fourier trigonométrica. Entonces el máximo de s_n se encuentra en $\pi/2n$ y el mínimo en $-\pi/2n$, y además,

$$\lim_{n\to\infty} s_n\left(\frac{\pi}{2n}\right) = \left(\frac{b-a}{2}\right)\left(\frac{2}{\pi}\int_0^\pi \frac{\operatorname{sen} t}{t}\, dt + 1\right) + b \approx (b - a)(0.089) + b.$$

Análogamente,

$$\lim_{n\to\infty} s_n\left(-\frac{\pi}{2n}\right) = \left(\frac{b-a}{2}\right)\left(-\frac{2}{\pi}\int_0^\pi \frac{\operatorname{sen} t}{t}\, dt + 1\right) + a \approx a - (b - a)(0.089),$$

y la diferencia de estos límites es

$$\left(\frac{b-a}{2}\right)\left(-\frac{4}{\pi}\int_0^\pi \frac{\operatorname{sen} t}{t}\, dt + 1\right) \approx (b - a)(1.179).$$

Demostración Demostraremos en primer lugar el caso particular $a = -1$, $b = 1$. Hemos visto que la serie de Fourier de

$$g(x) = \begin{cases} -1 & -\pi \leq x < 0 \\ 1 & 0 \leq x \leq \pi \end{cases}$$

es

$$\frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\operatorname{sen}[(2n-1)x]}{2n-1}.$$

Sea

$$s_n(x) = \frac{4}{\pi}\sum_{k=1}^{n}\frac{\operatorname{sen}[(2k-1)x]}{2k-1}.$$

Derivando, vemos que s_n tiene un máximo en $x_n = \pi/2n$ (algunos detalles los dejaremos al lector). El valor en ese máximo es

$$s_n\left(\frac{\pi}{2n}\right) = \frac{4}{\pi}\sum_{k=1}^{n}\frac{\operatorname{sen}[(2k-1)\pi/2n]}{2k-1} = \frac{2}{\pi}\sum_{k=1}^{n}\frac{\operatorname{sen}[(2k-1)\pi/2n]}{(2k-1)\pi/2n}\left(\frac{\pi}{n}\right).$$

Ésta es una suma de Riemann para la función $(\operatorname{sen} y)/y$ en $[0, \pi]$ para la partición $\{0, \pi/n, 2\pi/n, \ldots, \pi\}$. En consecuencia, si elegimos n par y hacemos $n \to \infty$, esto converge (véase el capítulo 8) a

$$\frac{2}{\pi}\int_0^{\pi}\frac{\operatorname{sen} y}{y}\,dy.$$

El caso del mínimo de f para $x < 0$ también se cumple, pues f y s_n son ambas impares.

El valor numérico de la integral es aproximadamente 1.179 y se calcula mediante métodos numéricos, como podrán verificar los lectores con sus calculadoras o computadores personales.

El caso de f general es consecuencia de la observación de que su serie de Fourier tiene como n-ésima suma parcial $(1/2)(b-a)(s_n+1)+a$ (¿por qué?). ■

10.6.1 Teorema de convergencia uniforme *Supóngase que f es continua en $[-\pi, \pi]$, $f(-\pi) = f(\pi)$ y que f' es continua a trozos, con discontinuidades de salto. Entonces la serie de Fourier (trigonométrica o exponencial) de f converge a f absoluta y uniformemente. Una afirmación similar se cumple para f en $[0, 2\pi]$.*

Demostración Podemos escribir

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \operatorname{sen} nx),$$

por el teorema **10.3.2**. Los coeficientes de Fourier de f' son

$$\alpha_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f'(x)\cos nx\,dx, \quad \beta_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f'(x)\operatorname{sen} nx\,dx,$$

y tenemos que

$$\alpha_n = \frac{n}{\pi}f(x)\cos nx\Big|_{-\pi}^{\pi} + \frac{n}{\pi}\int_{-\pi}^{\pi} f(x)\operatorname{sen} nx\,dx = nb_n,$$

ya que podemos integrar por partes y $f(\pi) = f(-\pi)$. Análogamente, $\beta_n = -na_n$.

Hay que tener cuidado con la justificación de la integración por partes, ya que f' existe solamente a trozos. Pero si la integración por partes se aplica en cada trozo usando el hecho de que f' tiene solamente discontinuidades de salto y observando la continuidad de f, entonces obtenemos los resultados.

Lema 13 *Bajo las condiciones del teorema* **10.6.1**, *tenemos*

$$\sum_{n=1}^{\infty} |a_n| < \infty, \quad \sum_{n=1}^{\infty} |b_n| < \infty,$$

y $na_n \to 0$, $nb_n \to 0$.

Demostración Sabemos que $\sum_{n=1}^{\infty}\beta_n^2$ converge (por la desigualdad de Bessel para f'). Sea $s_n = \sum_{k=1}^{n}|a_k|$,

$$s_n = \sum_{k=1}^{n}\frac{|\beta_k|}{k} = \sum_{k=1}^{n}\sqrt{\frac{\beta_k^2}{k^2}} \le \left\{\left(\sum_{k=1}^{\infty}\beta_k^2\right)\left(\sum_{k=1}^{\infty}1/k^2\right)\right\}^{1/2},$$

por la desigualdad de Schwarz. Como esto está acotado, también lo está s_n y por lo tanto converge (una sucesión creciente converge sii está acotada). Así, $\sum_{n=1}^{\infty}|a_n|$ converge. Como $\beta_n \to 0$, también tenemos que $na_n \to 0$. El caso de los términos b_n es similar. ■

Basta mostrar que $a_0/2 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \operatorname{sen} nx)$ converge uniformemente, ya que el límite debe ser $f(x)$. A su vez, basta mostrar que $\sum_{n=1}^{\infty}(a_n \cos nx + b_n \operatorname{sen} nx)$ es uniformemente convergente. Sin embargo,

$$|a_n \cos nx + b_n \operatorname{sen} nx| \le |a_n| + |b_n| = M_n,$$

y entonces, por el lema, $\sum_{n=1}^{\infty}M_n$ converge. Por lo tanto, por el criterio M de Weierstrass, la serie converge uniforme y absolutamente. ■

10.6.2 Teorema de derivación *Sea f continua en* $[-\pi, \pi]$, $f(-\pi) = f(\pi)$ *y* f' *continua a trozos, con discontinuidades de salto. Supóngase que* f'' *existe en* $x \in [-\pi, \pi]$. *Entonces la serie de Fourier para*

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \operatorname{sen} nx)$$

se puede derivar término a término en x para dar

$$f'(x) = \sum_{n=1}^{\infty}(-na_n \operatorname{sen} nx + nb_n \cos nx).$$

Además, ésta es la serie de Fourier de f'.

Demostración La demostración del teorema **10.6.1** probaba que los coeficientes de Fourier de f' están dados por

$$\alpha_n = nb_n, \quad \beta_n = -na_n.$$

Esta observación basta para demostrar el teorema, ya que si f'' existe, $f'(x)$ será la suma de la serie de Fourier. ■

10.7.1 Teorema *En el problema del desplazamiento inicial, supóngase que f es dos veces derivable. Entonces la solución a dicho problema es*

$$y(x,t) = \frac{1}{2}[f(x - ct) + f(x + ct)] = \sum_{n=1}^{\infty} b_n \operatorname{sen}\frac{n\pi x}{l} \cos\frac{n\pi ct}{l},$$

donde los b_n *son los coeficientes de la serie de senos en un semiintervalo,*

$$b_n = \frac{2}{l}\int_0^l f(x) \operatorname{sen}\frac{n\pi x}{l}\, dx,$$

y f se extiende de modo que sea impar y periódica (dos veces derivable significa que la función f extendida es dos veces derivable; véase la figura 10.7-2).

Demostración La serie de $y(x, t)$ converge porque $\sum_{n=1}^{\infty} b_n \operatorname{sen}(n\pi x/l)$ converge uniforme y absolutamente a f (teorema **10.6.1**). Vamos a mostrar que

$$\sum_{n=1}^{\infty} b_n \operatorname{sen}\frac{n\pi x}{l} \cos\frac{n\pi ct}{l} = \frac{1}{2}[f(x - ct) + f(x + ct)].$$

Para ello, obsérvese que

$$2 \operatorname{sen} \frac{n\pi x}{l} \cos \frac{n\pi ct}{l} = \operatorname{sen} \frac{n\pi(x - ct)}{l} + \operatorname{sen} \frac{n\pi(x + ct)}{l},$$

de modo que

$$\sum_{n=1}^{\infty} b_n \operatorname{sen} \frac{n\pi x}{l} \cos \frac{n\pi ct}{l} = \frac{1}{2} \sum_{n=1}^{\infty} b_n \operatorname{sen} \frac{n\pi(x - ct)}{l} + \sum_{n=1}^{\infty} b_n \operatorname{sen} \frac{n\pi(x + ct)}{l}$$
$$= \frac{1}{2}[f(x - ct) + f(x + ct)].$$

Ahora verificaremos que

$$y(x, t) = \frac{1}{2}[f(x - ct) + f(x + ct)]$$

satisface todas las condiciones. En primer lugar,

$$\frac{\partial^2 y}{\partial t^2} = \frac{1}{2} c^2 [f''(x - ct) + f''(x + ct)] = c^2 \frac{\partial^2 y}{\partial x^2}.$$

En segundo lugar, en $t = 0$, $y(x, 0) = f(x)$ y

$$\frac{\partial y}{\partial t}(x, 0) = \frac{1}{2} c[-f'(x) + f'(x) + f'(x)] = 0.$$

En tercer lugar, $y(0, t) = \frac{1}{2}[f(-ct) + f(ct)] = 0$, pues f es impar (al extenderla) y

$$y(l, t) = \frac{1}{2}[f(l - ct) + f(l + ct)] = 0,$$

pues $f(l - ct) = -f(ct - l) = -f(ct + l)$, ya que $f(x) = f(x + 2l)$ por la periodicidad. ■

10.7.2 Teorema *Si f es de cuadrado integrable, entonces, para cada $t > 0$,*

$$T(x, t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n e^{-n^2\pi^2 t/l^2} \cos \frac{n\pi x}{l}$$

converge uniformemente, es derivable y satisface la ecuación del calor y las condiciones de contorno. En $t = 0$ es igual a f en el sentido de la convergencia en media, y puntualmente si f es de clase C^1. Como siempre,

$$a_n = \frac{2}{l} \int_0^l f(x) \cos \frac{n\pi x}{l}\, dx.$$

Demostración Para mostrar que $T(x, t)$ satisface la ecuación del calor, lo que debemos hacer es justificar la derivación término a término con respecto a x y a t. Para ello, usamos el teorema **5.4.3**. Lo que debemos mostrar es que la serie de las derivadas

$$-\sum_{n=1}^{\infty} \frac{a_n \pi^2 n^2}{l^2} e^{-n^2\pi^2 t/l^2} \cos \frac{n\pi x}{l}$$

(que representa a $\partial T/\partial t$ y a $\partial^2 T/\partial x^2$) converge uniformemente en t y en x, lo que hacemos por medio del criterio M de Weierstrass. Como $|a_n|$ está acotada (de hecho $a_n \to 0$), podemos omitir los términos $a_n\pi^2/l^2$. Ahora, con respecto a x, sea $M_n = n^2 e^{-n^2\pi^2 t/l^2}$. Por el criterio del cociente, $\sum M_n < \infty$, de modo que la serie converge uniformemente con respecto a x.

Uniformemente en t significa uniformemente con respecto a todo $t \geq \varepsilon$, donde $\varepsilon > 0$ es arbitrario pero fijo. En ese caso, sea $M_n = n^2 e^{-n^2\pi^2 \varepsilon/l^2}$ y obsérvese que $\sum M_n$ converge (no podemos permitir que $t = 0$). El resto del teorema es obvio. ■

10.7.3 Teorema *En el teorema* **10.7.2**,

$$\lim_{t \to 0, t>0} T(x, t) = f(x)$$

en el sentido de la convergencia en media, y la convergencia es uniforme (y puntual) si f es continua, con f' continua a trozos. De modo más general, si la serie de Fourier de f converge en x a $f(x)$, entonces $T(x, t) \to f(x)$ cuando $t \to 0$.

Demostración Para la primera parte, bastará mostrar lo siguiente.

Lema 14 *Para cada $t > 0$, supóngase que $f_t \in V$, un espacio con producto interno y que $\varphi_0, \varphi_1, \ldots$ es una base ortonormal completa. Sea*

$$f_t = \sum_{n=1}^{\infty} c_n(t)\varphi_n, \quad f = \sum_{n=1}^{\infty} c_n\varphi_n.$$

Si

$$\lim_{t \to 0} \sum_{n=1}^{\infty} |c_n(t) - c_n|^2 = 0,$$

entonces $f_t \to f$ (en media).

Demostración El resultado es consecuencia de la relación de Parseval $\|f_t - f\|^2 = \sum_{n=1}^{\infty} |c_n(t) - c_n|^2$. ■

En el caso del teorema **10.7.3**, debemos mostrar que

$$\lim_{t\to 0}\sum_{n=1}^{\infty} |a_n|^2(1 - e^{-n^2\pi^2 t/l^2})^2 = 0.$$

Para hacerlo, basta mostrar que la función $g(t) = \sum_{n=1}^{\infty} |a_n|^2 (1 - e^{-n^2\pi^2 t/l^2})^2$ es continua en t, pues $g(0) = 0$. Para mostrar que $g(t)$ es continua, mostraremos que la serie converge uniformemente en t. Para ello usaremos el criterio de Abel. La forma que necesitamos es la siguiente:

Lema 15 *Sea* $\sum_{n=1}^{\infty} c_n$ *una serie convergente y* $\varphi_n(t)$ *una sucesión uniformemente acotada, decreciente (respectivamente, creciente) definida para* $t \geq 0$. *Entonces* $g(t) = \sum_{n=1}^{\infty} c_n\varphi_n(t)$ *converge uniformemente en t. En particular, g es continua y* $g(0) = \lim_{t\to 0} g(t)$.

Véase el teorema **5.9.1** para la demostración de este lema. El caso creciente se deduce del caso decreciente, analizando $-g(t)$ en vez de $g(t)$. En nuestro caso, $c_n = |a_n|^2$ y $\varphi_n(t) = (1 - e^{-n^2\pi^2 t/l^2})^2$. Ahora, $\varphi_n \leq \varphi_m$ si $n \leq m$ y $|\varphi_n(t)| \leq 1$. Así, por el lema y el hecho de que $\sum c_n$ converge, tenemos nuestro resultado.

Supóngase ahora que f' es continua a trozos. Por la demostración del teorema **10.6.1**, $\sum_{n=1}^{\infty} |a_n| < \infty$. Así, para x dado,

$$|f(x) - T(x,t)| \leq \sum_{n=1}^{\infty} |a_n|(1 - e^{-n^2\pi^2 t/l^2}).$$

Por un argumento como el anterior, la serie de la derecha converge uniformemente, de modo que podemos hacer $t \to 0$ en cada término para concluir que $T(x, t) \to f(x)$ cuando $t \to 0$. De hecho, obsérvese que la convergencia es uniforme en x, pues tenemos la cota $\sum_{n=1}^{\infty} |a_n| (1 - e^{-n^2\pi^2 t/l^2})$, que tiende a 0 cuando $t \to 0$ y es independiente de x.

Por último, supóngase que $\sum_{n=1}^{\infty} a_n \cos(n\pi x/l)$ converge para algún x fijo. Entonces queremos mostrar que (para este x fijo)

$$\lim_{t\to 0} g(t) = \lim_{t\to 0}\sum_{n=1}^{\infty} a_n e^{-n^2\pi^2 t/l}\cos\frac{n\pi x}{l} = 0.$$

Aquí no podemos hacer la misma estimación, pues el factor $\cos(n\pi x/l)$ es necesario para la convergencia de $\sum a_n \cos(n\pi x/l)$. Sin embargo, podemos aplicar el lema **15** con $c_n = a_n \cos(n\pi x/l)$ y $\varphi_n(t) = e^{-n^2\pi^2 t/l^2}$ para obtener la conclusión, ya que las φ_n son decrecientes y están acotadas por 1. ■

También podemos concluir de esta demostración que

$$\lim_{t\to t_0} T(x,t) = T(x,t_0);$$

es decir, T es continua en t en cada uno de los tres casos del teorema **10.7.3**. De hecho, ya sabemos que para $t > 0$, $T(x, t)$ es derivable, y por tanto continua. Sin embargo, $T(x, t)$ podría no ser derivable en $t = 0$, pero la demostración anterior muestra que sí tenemos continuidad en $t = 0$.

Estos métodos que usan los criterios de Abel y de Dirichlet son importantes para establecer la convergencia en otros problemas (como la ecuación de Laplace), como veremos en la siguiente demostración.

10.7.4 Teorema

i. *Dado g_1, sea $\varphi(x, y)$ dada por*

$$\varphi(x,y) = \sum_{n=1}^{\infty} b_n \operatorname{senh} \frac{n\pi(b-y)}{a} \frac{\operatorname{sen}(n\pi x/a)}{\operatorname{senh}(n\pi b/a)}, \tag{12}$$

Supóngase que g_1 es de clase C^2 y que $g_1(0) = g_1(a) = 0$. Entonces φ converge uniformemente; es la solución al problema de Dirichlet con $f_1 = f_2 = g_2 = 0$; es continua en todo el cuadrado y $\nabla^2\phi = 0$ en el interior.

ii. *Si cada f_1, f_2, g_1, g_2 es de clase C^2 y se anula en las esquinas del rectángulo, entonces la solución $\varphi(x, y)$ es la suma de cuatro series como la de la ecuación (12), $\nabla^2\varphi = 0$ en el interior; φ es continua en todo el rectángulo y alcanza los valores dados sobre el contorno. Además, φ es C^∞ en el interior.*

iii. *Si f_1, f_2, g_1, g_2 sólo son de cuadrado integrable, entonces la serie de φ converge en el interior, $\nabla^2\varphi = 0$ y φ es C^∞. Además, φ alcanza los valores de contorno en el sentido de la convergencia en media. Esto significa, por ejemplo, que $\lim_{y\to 0} \varphi(x, y) = \varphi(x, 0)$ $g_1(x)$ con la convergencia en media.*

Demostración Para simplificar, analizaremos el caso $a = b = \pi$; el caso general se obtiene mediante un cambio de coordenadas. Para demostrar las partes **i** y **ii** del teorema, mostraremos que $\varphi(x,y)$ converge uniformemente en x e y, y que podemos derivar dos veces, término a término, en el interior. En vista de las observaciones anteriores, esto basta para demostrar el teorema. La parte **ii** es una consecuencia de **i** y de la linealidad: los valores en la frontera son válidos, pues g_1 está representada por su serie de Fourier.

Por los teoremas **10.6.1** y **10.6.2**,

$$g_1(x) = \sum_{n=1}^{\infty} b_n \operatorname{sen} nx, \quad g_1'(x) = \sum_{n=1}^{\infty} nb_n \cos nx,$$

y estas series convergen uniforme y absolutamente. Aquí usamos el hecho de que $g_1(0) = g_1(\pi) = 0$.

Para mostrar que φ converge uniformemente usamos el criterio de Abel, como en la demostración del teorema **10.7.3**, en el cuadrado $[0, \pi] \times [0, \pi]$. Así, debido a que la serie de g_1 converge uniformemente, debemos mostrar que $\varphi_n = (\operatorname{sen} n(\pi - y))/\operatorname{sen} n\pi$ es decreciente con n y que estas funciones están uniformemente acotadas. Si podemos demostrar que son decrecientes, eso implica que están uniformemente acotadas, pues $0 \le \varphi_n(y) \le \varphi_1(y)$ y φ_1 está acotada, ya que es continua. De hecho, en este caso, $\varphi_1 \le 1$.

Para mostrar que $\varphi_{n+1} \le \varphi_n$, fijamos y y consideramos la función $\psi(t) = (\operatorname{senh} t(\pi - y))/\operatorname{senh} t\pi$, $t > 0$. Basta mostrar que $\psi'(t) \le 0$, ya que entonces ψ decrece cuando t crece y, en particular, $\psi(n + 1) \le \psi(n)$. Éste es un caso particular del siguiente lema.

Lema 16 *Para las constantes α, β que satisfacen $\beta > 0$, $\beta \ge \alpha$ y si $\psi(t) = \operatorname{senh}(\alpha t)/\operatorname{senh}(\beta t)$, $\psi'(t) \le 0$ para $t \ge 0$.*

Demostración

$$\begin{aligned}\operatorname{senh}^2(\beta t)\,\psi'(t) &= \alpha\operatorname{senh}(\beta t)\cosh(\alpha t) - \beta\operatorname{senh}(\alpha t)\cosh(\beta t)\\ &= -\frac{\beta^2-\alpha^2}{2}\left[\frac{\operatorname{senh}[(\alpha+\beta)t]}{\alpha+\beta} - \frac{\operatorname{senh}[(\beta-\alpha)t]}{\beta-\alpha}\right],\end{aligned}$$

usando la identidad $\operatorname{senh}(u + v) = \operatorname{senh} u \cosh v + \operatorname{senh} v \cosh u$. Si el término entre corchetes es ≥ 0, hemos terminado, pues $\beta^2 - \alpha^2 > 0$. Esto es, de hecho, cierto. Para verlo, sea

$$\rho(t) = \frac{\operatorname{senh}[(\alpha+\beta)t]}{\alpha+\beta} - \frac{\operatorname{senh}[(\beta-\alpha)t]}{\beta-\alpha}.$$

Ahora, $\rho(0) = 0$ y $\rho'(t) = \operatorname{senh}[(\alpha + \beta)t] - \operatorname{senh}[(\beta - \alpha)t] \ge 0$, pues senh es creciente. Por lo tanto, $\rho(t) \ge 0$ para todo $t \ge 0$. ■

Esto establece la primera parte de la demostración, que dice que la serie de $\varphi(x, y)$ converge uniformemente. Para la parte de la derivabilidad, debemos mostrar que

$$\lambda(x, y) = \sum_{n=1}^{\infty} n^2 b_n \operatorname{senh} n(\pi - y)\,\frac{\operatorname{sen} nx}{\operatorname{senh}(n\pi)}$$

converge uniformemente (ésta es la segunda derivada formal con respecto a y; la segunda derivada con respecto a x es su negativa).

Es importante observar que podemos obtener la convergencia uniforme sólo si podemos estar lejos de la frontera; de hecho, para cualquier $\varepsilon > 0$ estableceremos la con-

vergencia uniforme en $0 < \varepsilon \le y \le \pi$ para x arbitrario. Con esta restricción adicional, ya no necesitamos la fineza del criterio de Abel; bastará con el criterio M de Weierstrass. Tenemos $|b_n| \le M$. Sea

$$M_n = n^2 M \frac{\operatorname{senh}[n(\pi - \varepsilon)]}{\operatorname{senh}(\pi n)}.$$

Entonces M_n acota los términos en λ. Pero $2\operatorname{senh}[n(\pi - \varepsilon)] < e^{n(\pi-\varepsilon)}$ y $2\operatorname{senh}(n\pi) \ge e^{n\pi}(1 - e^{n\pi})$, por la definición de senh. Así,

$$M_n \le Mn^2 \frac{e^{-n\varepsilon}}{1 - e^{-2\pi}}.$$

Como $\varepsilon > 0$, $\sum M_n$ converge, por lo que tenemos convergencia uniforme.

Obsérvese que podríamos usar n^k en vez de n^2 para cualquier k y seguir teniendo la convergencia; de hecho, podemos derivar cualquier número de veces; es decir, φ es C^∞ (un poco de reflexión muestra que φ es analítica; véase el ejemplo **6.8.7**).

La demostración de la parte **iii** es rutinaria. Para mostrar que $\nabla^2\varphi = 0$ y que φ es una función C^∞ en el interior, la demostración es similar a la anterior (lo único que hemos usado es que los b_n están acotados). Para la convergencia en media, procedemos como en la demostración del teorema **10.7.3**, usando el lema **15.**

Ejemplos resueltos del capítulo 10

Ejemplo 10.1 Sea f: $[0, \pi] \to \mathbb{C}$ una función continua. Demuéstrese que es válida la siguiente desigualdad:

$$\left|\int_0^\pi f(x) \operatorname{sen} x\, dx\right|^2 + \cdots + \left|\int_0^\pi f(x) \operatorname{sen} nx\, dx\right|^2 \le \frac{\pi}{2}\int_0^\pi |f(x)|^2\, dx.$$

Solución Esto es consecuencia de la desigualdad de Bessel aplicada al siguiente sistema ortonormal (incompleto) en $[0, \pi]$:

$$\sqrt{2/\pi}\, \operatorname{sen} x, \ldots, \sqrt{2/\pi}\, \operatorname{sen} nx.$$

Si usamos una suma infinita tendríamos la igualdad, por el teorema de Parseval (véase la tabla 10.5-3). ♦

Ejemplo 10.2 *Muéstrese que si $f_n \to f$ (en media) y $g_n \to g$ (en media) en un espacio con producto interno, entonces*

$$\langle f_n, g_n\rangle \to \langle f, g\rangle.$$

Solución En primer lugar, haremos una estimación por medio de la desigualdad de Schwarz y la desigualdad triangular:

$$\begin{aligned}|\langle f_n, g_n\rangle - \langle f, g\rangle| &\le |\langle f_n, g_n\rangle - \langle f_n, g\rangle| + |\langle f_n, g\rangle - \langle f, g\rangle| \\ &= |\langle f_n, g_n - g\rangle| + |\langle f_n - f, g\rangle| \\ &\le \|f_n\|\,\|g_n - g\| + \|f_n - f\|\,\|g\|.\end{aligned}$$

El resultado es consecuencia de esto. Dado $\varepsilon > 0$, elegimos N de modo que $n \ge N$ implique cada una de las siguientes estimaciones:

1. $\|f_n - f\| < \frac{\varepsilon}{2}\|g\|$;
2. $\|f_n - f\| \le 1$; y
3. $\|g_n - g\| < \dfrac{\varepsilon}{2(\|f\| + 1)}$.

(¿Por qué es posible esto?)

Entonces $\|f_n - f\| \le 1$ implica $\|f_n\| \le \|f\| + 1$ (¿por qué?), y para $n \ge N$,

$$\begin{aligned}|\langle f_n, g_n\rangle - \langle f, g\rangle| &\le (\|f\| + 1)\|g_n - g\| + \|f_n - f\|\,\|g\| \\ &\le (\|f\| + 1)\frac{\varepsilon}{2}(\|f\| + 1)^{-1} + \frac{\varepsilon\|g\|}{2\|g\|} = \varepsilon.\end{aligned}$$

Esto demuestra que $\langle f_n, g_n\rangle \to \langle f, g\rangle$. ♦

Ejemplo 10.3 *Sea* $f: [0, 2\pi] \to \mathbb{R}$ *de cuadrado integrable y defínase* $g(x) = \int_0^x f(y)\,dy$. *Determínese el desarrollo de Fourier de g y establézcase dónde es válido.*

Solución Como hicimos en el teorema **10.5.1**, podemos integrar la serie de Fourier

$$f(y) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos ny + b_n \operatorname{sen} ny)$$

término a término desde $y = 0$ hasta $y = x$ para obtener

$$g(x) = \int_0^x f(y)\,dy = \frac{a_0}{2}x + \sum_{n=1}^{\infty}\left(\frac{a_n}{n}\operatorname{sen} nx + \frac{b_n}{n}(1 - \cos nx)\right).$$

Por la tabla 10.5-4 sabemos que $x = \pi - 2\sum_{n=1}^{\infty}(1/n)\operatorname{sen} nx$ para x en $[0, 2\pi]$, convergente en media y puntualmente para x distinto de 0 o 2π. Además, la desigualdad de

Cauchy-Schwarz muestra que $\sum(b_n/n)$ converge, pues tanto $\sum b_n^2$ como $\sum(1/n)^2$ lo hacen. Si juntamos estos resultados, podemos reescribir la última serie como una serie de Fourier:

$$g(z) = \frac{a_0\pi}{2} + \sum_{k=1}^{\infty} \frac{b_n}{n} + \sum_{k=1}^{\infty} \left(\frac{a_n - a_0}{n} \operatorname{sen} nx - \frac{b_n}{n} \cos nx \right).$$

que converge en media y puntualmente para x distinto de 0 o 2π. Puesto que los coeficientes de Fourier son únicos, éste es el desarrollo de $g(x)$.

Como g está acotada por $\int_0^{2\pi} |f(x)|\, dx$, que es finita, g es de cuadrado integrable y, por lo tanto, tiene una serie de Fourier. Obsérvese que g es continua pero no necesariamente derivable (véase el ejercicio 61). ♦

Ejemplo 10.4 *Sean $f : [0, 2\pi] \to \mathbb{R}$ y $g : [0, 2\pi] \to \mathbb{R}$ extendidas periódicamente. Defínase la **convolución** de f y g como*

$$(f * g)(x) = \int_0^{2\pi} f(y)g(x-y)\, dy.$$

Calcúlese la serie de Fourier de $f * g$ en términos de las series de f y de g, usando la forma exponencial de las series de Fourier.

Solución Los coeficientes de Fourier de $f * g$ están dados por

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} (f * g)(x) e^{-inx}\, dx = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} f(y)g(x-y)e^{-inx}\, dy\, dx$$
$$= \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} f(y)e^{-iny} g(x-y)e^{-in(x-y)}\, dy\, dx,$$

pues $e^{-inx} = e^{-iny} \cdot e^{-inx+iny}$. Si cambiamos de variables ($y \mapsto y$ y $x - y \mapsto t$) y usamos la periodicidad (podemos intercambiar el orden de integración gracias al teorema de Fubini, §9.2), obtenemos

$$\frac{1}{2\pi} \int_0^{2\pi} f(y)e^{-iny}\, dy \int_0^{2\pi} g(t)e^{-int}\, dt.$$

Sean $f(x) = \sum_{n=-\infty}^{\infty} a_n e^{inx}$ y $g(x) = \sum_{n=-\infty}^{\infty} b_n e^{inx}$ las series de Fourier de f y g, de modo que

$$a_n = \frac{1}{2\pi} \int_0^{2\pi} f(x)e^{-inx}\, dx \quad \text{y} \quad b_n = \frac{1}{2\pi} \int_0^{2\pi} g(x)e^{-inx}\, dx.$$

Entonces, el cálculo anterior muestra que tenemos

$$c_n = 2\pi a_n b_n, \text{ con lo que } (f * g)(x) = 2\pi \sum_{-\infty}^{\infty} a_n b_n e^{inx}.$$

Podemos hacer una operación similar con las series trigonométricas, aunque los cálculos y resultados son más complicados. Para que los cálculos anteriores sean válidos se necesitan condiciones suficientes sobre f y g, como el hecho de que sean continuas a trozos o, más generalmente, de cuadrado integrables. Si queremos que la serie converja puntualmente, debemos agregar la hipótesis del teorema de la convergencia puntual **10.3.2.** ♦

Ejemplo 10.5 *Considérese $f(x) = \cos \lambda x$, $-\pi \leq x \leq \pi$, donde λ es real y no entero. Calcúlese la serie de Fourier de f.*

Solución

$$\begin{aligned} c_n &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos \lambda x e^{-inx}\,dx = \frac{1}{4\pi}\int_{-\pi}^{\pi} (e^{i\lambda x} + e^{-i\lambda x})e^{-inx}\,dx \\ &= \frac{1}{4\pi}\left[\frac{e^{i(\lambda-n)x}}{i(\lambda-n)} + \frac{e^{-i(\lambda+n)x}}{-i(\lambda+n)}\right]_{x=-\pi}^{\pi} \\ &= \frac{(-1)^n}{4\pi i}\cdot 2i \operatorname{sen}\lambda\pi \left(\frac{1}{\lambda-n} + \frac{1}{\lambda+n}\right) = \frac{(-1)^n}{2\pi}\operatorname{sen}\lambda\pi\,\frac{2\lambda}{\lambda^2-n^2}.\end{aligned}$$

Por el teorema de la convergencia puntual, la serie de Fourier converge a $f(x)$ en todos los puntos. Por lo tanto, para $|x| \leq \pi$,

$$\cos \lambda x = \frac{\operatorname{sen}\pi\lambda}{2\pi}\sum_{k=-\infty}^{\infty}(-1)^k\frac{2\lambda}{\lambda^2-k^2}e^{ikx}.$$

Formalmente, la solución está completa, pero continuemos. Si $x = \pi$, entonces

$$\cos\pi\lambda = \frac{\operatorname{sen}\pi\lambda}{2\pi}\sum_{k=-\infty}^{\infty}(-1)^k\frac{2\lambda}{\lambda^2-k^2}(-1)^k = \frac{\operatorname{sen}\pi\lambda}{\pi}\left\{\frac{1}{\lambda}+\sum_{k=1}^{\infty}\frac{2\lambda}{\lambda^2-k^2}\right\}.$$

En consecuencia,

$$\pi\frac{\cos\pi\lambda}{\operatorname{sen}\pi\lambda} - \frac{1}{\lambda} = \sum_{k=1}^{\infty}\frac{2\lambda}{\lambda^2-k^2}, \quad \lambda \neq \text{entero}.$$

Obsérvese que la serie de la derecha converge uniformemente para $0 \leq \lambda \leq \lambda_0 < 1$. Obsérvese también que $\pi(\cos\pi\lambda/\operatorname{sen}\pi\lambda) - (1/\lambda) \to 0$ cuando $\lambda \to 0$, por lo que es integrable Riemann. Integrando,

$$\log\left(\frac{\operatorname{sen}\pi\lambda}{\pi\lambda}\right) = \sum_{k=1}^{\infty}\log\left(1-\frac{\lambda^2}{k^2}\right), \quad |\lambda| < 1.$$

Tomando exponenciales,

$$\frac{\operatorname{sen} \pi\lambda}{\pi\lambda} = \prod_{k=1}^{\infty}\left(1 - \frac{\lambda^2}{k^2}\right); \quad \text{es decir,} \quad \operatorname{sen} \pi\lambda = \pi\lambda \prod_{k=1}^{\infty}\left(1 - \frac{\lambda^2}{k^2}\right), \quad |\lambda| < 1.$$

El producto del miembro derecho define una función de λ de periodo 2 (véase el ejercicio 74) al igual que el miembro izquierdo, por lo que esta fórmula es válida para *todo* valor real de λ. Esta fórmula del producto para el seno fue descubierta (aunque no fue demostrada de forma rigurosa) por Euler (para ver otro método de demostración, consúltese J. Marsden y M. Hoffman, *Basic Complex Analysis, op. cit.*, capítulo 7).

Si $\lambda = 1/2$, entonces

$$1 = \frac{\pi}{2}\prod_{k=1}^{\infty}\left(1 - \frac{1}{4k^2}\right) = \frac{\pi}{2}\prod_{k=1}^{\infty}\frac{(2k-1)(2k+1)}{(2k)\,(2k)},$$

por lo que

$$\frac{\pi}{2} = \prod_{k=1}^{\infty}\frac{(2k)\,(2k)}{(2k-1)(2k+1)}; \quad \text{es decir,}$$

$$\frac{\pi}{2} = \lim_{n\to\infty}\frac{(2\cdot 2)(4\cdot 4)(6\cdot 6)\cdots(2n\cdot 2n)}{(1\cdot 3)(3\cdot 5)(5\cdot 7)\cdots((2n-1)\cdot(2n+1))},$$

que recibe el nombre de ***fórmula del producto de Wallis*** para $\pi/2$. ♦

Ejemplo 10.6 *Una aplicación interesante de la relación de Parseval al problema isoperimétrico es la siguiente: muéstrese que de entre todas las curvas planas con un perímetro dado, el área máxima queda encerrada por una circunferencia.*

Solución Sea $(x(t), y(t))$, $0 \le t \le 2\pi$ una representación paramétrica de una curva simple cerrada. Suponemos que $x(t)$, $y(t)$ son funciones C^1 de t y que el parámetro t es la longitud de arco. Así, la longitud total es 2π y $\dot{x}(t)^2 + \dot{y}(t)^2 = 1$, de modo que $\int_0^{2\pi}(\dot{x}(t)^2 + \dot{y}(t)^2)dt = 2\pi$. El área encerrada es

$$A = \int_0^{2\pi} x(t)\dot{y}(t)\,dt.$$

(Ésta es una fórmula conocida del cálculo; véase, por ejemplo, Marsden y Tromba, *Cálculo vectorial, op. cit.*, capítulo 7.) Afirmamos que $A \le \pi$, y que $A = \pi$ sólo si la curva es una circunferencia. Para demostrarlo, expresamos A en términos de los coeficientes de Fourier de x e y. Escribimos

$$x(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k \cos kt + b_k \operatorname{sen} kt)$$

y

$$y(t) = \frac{\alpha_0}{2} + \sum_{k=1}^{\infty}(\alpha_k \cos kt + \beta_k \operatorname{sen} kt).$$

Todos los coeficientes son reales, luego, por un cambio de origen en el plano, podemos suponer que $a_0 = \alpha_0 = 0$.

Las series de Fourier de las derivadas $\dot{x}(t)$, $\dot{y}(t)$ son

$$\dot{x}(t) = \sum_{k=1}^{\infty}(kb_k \cos kt + ba_k \operatorname{sen} kt)$$

y

$$\dot{y}(t) = \sum_{k=1}^{\infty}(k\beta_k \cos kt + b\alpha_k \operatorname{sen} kt).$$

Por la relación de Parseval,

$$2\pi = \int_0^{2\pi}(\dot{x}^2 + \dot{y}^2)\,dt = \pi \sum_{k=1}^{\infty} k^2(a_k^2 + b_k^2 + \alpha_k^2 + \beta_k^2)$$

y el área es

$$A = \int_0^{2\pi} x\dot{y}\,dt = \pi \sum_{k=1}^{\infty} k(a_k\beta_k - b_k\alpha_k).$$

Por lo tanto

$$\begin{aligned} 2\pi - 2A &= \pi \sum_{k=1}^{\infty}\{k^2(a_k^2 + b_k^2 + \alpha_k^2 + \beta_k^2) - 2k(a_k\beta_k - b_k\alpha_k)\} \\ &= \pi \sum_{k=1}^{\infty}(k^2 - k)(a_k^2 + b_k^2 + \alpha_k^2 + \beta_k^2) \\ &\quad + \pi \sum_{k=1}^{\infty} k\{(a_k - \beta_k)^2 + (\alpha_k + b_k)^2\}. \end{aligned}$$

Así, $\pi - A \geq 0$ y $\pi - A = 0$ si y sólo si

i. $a_k = b_k = \alpha_k = \beta_k = 0$ para $k \geq 2$;

ii. $a_1 = \beta_1$, $\alpha_1 = -b_1$.

En este caso,

$$x(t) = a_1 \cos t + b_1 \operatorname{sen} t \qquad \text{y} \qquad y(t) = -b_1 \cos t + a_1 \operatorname{sen} t = -x(t + \pi/2).$$

De forma equivalente, para algún R y δ,

$$x(t) = R\cos(t+\delta) \qquad \text{y} \qquad y(t) = R\,\text{sen}(t+\delta).$$

La condición $x^2 + y^2 = 1$ implica que $R = 1$, y por lo tanto tenemos una circunferencia de radio 1. ♦

Ejercicios del capítulo 10

1. Sea V un espacio con producto interno y $M \subset V$ un subespacio vectorial. Defínase el ***complemento ortogonal*** de M como

$$M^{\perp} = \{ f \in V \mid \langle f, g \rangle = 0 \text{ para todo } g \in M \}.$$

Muéstrese que $M^{\perp}$ es un subespacio vectorial de V y que es cerrado (es decir, si $f_n \in M^{\perp}$ y $f_n \to f$ (en media), entonces $f_n \in M^{\perp}$).

2. Demuéstrese que los polinomios de Legendre (véase **§10.2**) son completos en L^2 de $[-1, 1]$. [Sugerencia: muéstrese primero que cualquier polinomio se puede desarrollar en términos de polinomios de Legendre y utilícese después el método de demostración del teorema de completitud media.]

3. **a.** Úsese la serie de Fourier de e^x en $[-\pi, \pi]$ para demostrar la siguiente identidad:

$$\frac{1}{2}(\pi \coth \pi - 1) = \sum_{1}^{\infty} \frac{1}{n^2+1}.$$

b. Úsese la serie de cosenos en un semiintervalo para $\cos ax$, donde a no es un entero, para demostrar la siguiente identidad:

$$\pi \cot \pi a = \frac{1}{a} + \sum_{1}^{\infty} \frac{2a}{a^2 - n^2}.$$

4. Demuéstrese que si $f_n \to f$ (en media), entonces $\|f_n\| \to \|f\|$. ¿Es cierto el recíproco?

5. Demuéstrese que la convergencia uniforme implica la convergencia en media (en intervalos finitos).

6. Considérese el espacio l_2 de todas las sucesiones $x = (x_1, x_2, \ldots)$ de números reales tales que $\sum_{i=1}^{\infty} x_i^2 < \infty$. Muéstrese que l_2 es un espacio con producto interno con $\langle x, y \rangle = \sum_{i=1}^{\infty} x_i y_i$. Además, muéstrese que este espacio es completo (mostrando que todas las sucesiones de Cauchy convergen).

7. Sea

$$l_2 = \left\{ (x_1, x_2, \dots) \;\middle|\; \sum_{i=1}^{\infty} x_i^2 < \infty \right\},$$

que, por el ejercicio 6, es un espacio de Hilbert. Sean $\varphi_n \in l_2$, $n = 1, 2, \dots$, dados por $\varphi_n = (0, 0, \dots, 1, 0, \dots)$, con el 1 en el n-ésimo lugar. Muéstrese de dos formas que $\varphi_1, \varphi_2, \dots$ es un conjunto ortonormal completo.

a. Directamente.

b. Usando el teorema de Parseval (véase también el ejercicio 14c).

8. Encuéntrense funciones f_n y f en $[-1, 1]$ tales que $f_n \to f$ (puntualmente) pero no en media.

9. Verifíquese que todas las funciones continuas a trozos $f: [a, b] \to \mathbb{C}$ forman un espacio vectorial.

10. Encuéntrese una sucesión f_n de funciones continuas a trozos tales que $f_n \to f$ puntualmente (respectivamente, en media) pero que f no sea continua a trozos.

11. Demuéstrese que si $f: [a, b] \to \mathbb{C}$ es continua a trozos, entonces también lo es $|f|$. Muéstrese también que $|f|$ es acotada.

12. Demuéstrese que si f y g son de cuadrado integrable en $[a, b]$, entonces $\|f - g\| = 0$ sii $f = g$ excepto en un conjunto de medida nula.

13. En la demostración de la desigualdad de Bessel, mostramos que

$$\left\| f - \sum_{k=0}^{n} t_k \varphi_k \right\|^2 = \|f\|^2 - \sum_{k=0}^{n} |\langle f, \varphi_k \rangle|^2 + \sum_{k=0}^{n} |\langle f, \varphi_k \rangle - t_k|^2.$$

Úsese esta igualdad con $t_k = \langle f, \varphi_k \rangle$ para dar otra demostración de la desigualdad de Bessel.

14. Supóngase que V es un espacio de Hilbert (es decir, un espacio completo con producto interno). Sea $\varphi_0, \varphi_1, \dots$ un conjunto ortonormal en V.

a. Para cada $f \in V$, muéstrese que

$$s_n = \sum_{k=0}^{n} \langle f, \varphi_k \rangle \varphi_k$$

converge a algún elemento de V. Sugerencia: muéstrese que

$$\|s_n - s_m\|^2 = \sum_{i=n+1}^{m} |\langle f, \varphi_i \rangle|^2,$$

y úsese la desigualdad de Bessel para mostrar que s_n es una sucesión de Cauchy.

b. Para cada $f \in V$, muéstrese que, si $s = \sum_{i=0}^{\infty} \langle f, \varphi_i \rangle \varphi_i$, entonces $f - s$ es ortogonal a cada φ_i y $f - s$ es ortogonal a s.

c. Muéstrese que si siempre que f sea ortogonal a cada φ_i tenemos $f = 0$, entonces $\varphi_0, \varphi_1, \ldots$ es completo. [Sugerencia: por **b**, $f - s$ es ortogonal a cada φ_i, y entonces $f = s$.]

15. Muéstrese que en V del teorema **10.1.2** no tenemos el teorema de Bolzano-Weierstrass; es decir, en un conjunto cerrado y acotado, una sucesión no tiene por qué tener subsucesiones convergentes.

16. Calcúlese la serie de Fourier de $f(x) = (x + x^2)/2$, $0 \leq x \leq 2\pi$.

17. **a.** Sean $\varphi_0, \varphi_1, \ldots$ vectores ortonormales en un espacio con producto interno V. Si

$$\sum_{k=0}^{\infty} c_k \varphi_k = 0,$$

muéstrese entonces que $c_k = 0$, $k = 0, 1, 2, \ldots$.

b. Si $\varphi_0, \varphi_1, \ldots$ es un conjunto completo y $\langle f, \varphi_i \rangle = \langle g, \varphi_i \rangle$ para todo i, entonces demuéstrese que $f = g$.

c. Si V es un espacio de funciones integrables, demuéstrese que **b** implica $f(x) = g(x)$ excepto posiblemente en un conjunto de medida nula.

18. Éste es un ejercicio acerca de las series de Fourier en varias variables.

a. Hemos visto que $e^{inx}/\sqrt{2\pi}$, $n = 0, \pm 1, \ldots$, son funciones ortonormales en $[0, 2\pi]$. Considérense las funciones

$$\varphi_{n,m}(x, y) = \frac{e^{inx+imy}}{2\pi} = \frac{e^{inx}}{\sqrt{2\pi}} \frac{e^{imy}}{\sqrt{2\pi}}.$$

Muéstrese que $\varphi_{n,m}$ son ortonormales en $[0, 2\pi] \times [0, 2\pi]$.

b. Generalícese **a** para construir $\varphi_{n,m}$, dadas unas φ_n generales que sean ortonormales en $[a, b]$.

Nota. Las $e^{inx}/\sqrt{2\pi}$ son completas, por lo que también lo son las $\varphi_{n,m}$. Esto se demuestra en el teorema **10.3.1** (el teorema de completitud media) para $e^{inx}/\sqrt{2\pi}$, y la demostración para $\varphi_{n,m}$ es similar.

19. ***Problemas de Sturm-Liouville*** Considérese la ecuación diferencial

$$\frac{d^2f}{dx^2} + [q(x) + \lambda p(x)]\, f(x) = 0$$

en la que debe determinarse $f(x)$, con condiciones de contorno $f(a) = f(b) = 0$ y $a \leq x \leq b$. Las funciones q, p están fijas, y suponemos que $p(x) > 0$. Los valores λ para los que existen soluciones f son los ***autovalores.***

a. Muéstrese que si f y g son soluciones con autovalores λ y μ y $\lambda \neq \mu$, entonces

$$\int_a^b p(x)f(x)g(x)\,dx = 0.$$

Sugerencia: Úsese la ecuación diferencial para mostrar que

$$(\lambda - \mu)p(x)f(x)g(x) = \frac{d}{dx}[g'(x)f(x) - f'(x)g(x)],$$

e intégrese después.

b. Interprétese **a** como la ortogonalidad de f y g con

$$\langle f, g\rangle = \int_a^b p(x)f(x)g(x)\,dx.$$

Muéstrese que éste es un producto interno.[5]

20. Muéstrese que $\{(1/\sqrt{\pi})$ sen $nx \mid n = 1, 2, \ldots\})$ es una familia ortonormal en $[0, 2\pi]$. ¿Cuál sería una serie de Fourier para esta familia? ¿Es completa?

21. Sea $\varphi_0, \varphi_1, \ldots$ un sistema ortonormal en un espacio con producto interno V y sea $f \in V$. Sea $s_n \sum_{k=0}^{n} \langle f, \varphi_k\rangle \varphi_k$ la n-ésima suma parcial de la serie de Fourier. Muéstrese que para cualquier entero $p \geq 0$,

$$\|f - s_{n+p}\| \leq \|f - s_n\|.$$

22. Sea V un espacio con producto interno y $\varphi_0, \varphi_1, \ldots$ un sistema ortonormal completo. Para $f, g \in V$, muéstrese que

$$\langle f, g\rangle = \sum_{k=0}^{\infty} \langle f, \varphi_k\rangle \langle \varphi_k, g\rangle.$$

Por supuesto, parte de este problema es mostrar que la serie converge. [Sugerencia: escríbanse f y g como límites de series de Fourier y aplíquese el ejemplo resuelto **10.2.**]

[5] Muchos sistemas ortonormales surgen de esta forma. El sistema trigonométrico surge con $p = 1$, $q = 0$. Existe un teorema avanzado que afirma que tales sistemas son completos. Véase, por ejemplo, Coddington y Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, Nueva York.

Los ejercicios 23–28 se refieren a sistemas en mecánica cuántica (§**10.9**).

23. Muéstrese que la solución de $i\hbar(\partial\psi/\partial t) = H\psi$, si $\psi = \psi_0$ en $t = 0$, está dada por

$$\psi = \sum_{n=0} \langle \psi_0, \varphi_n \rangle e^{-iE_n t/\hbar} \varphi_n,$$

donde las $\varphi_n \in V$ son las autofunciones de H con autovalores E_n (puede suponerse que la serie se puede derivar término a término y que las autofunciones son completas). ¿Qué ocurre si ψ ya es una autofunción?

24. Calcúlense los conmutadores de los operadores Q_x, Q_y, Q_z, P_x, P_y, P_z y J_x, J_y, J_z, dados en el texto. ¿Qué dice el principio de incertidumbre para estos operadores (ejercicio 5, §**10.9**)?

26. Si A y B son simétricos, ¿es $[A, B]$ simétrico? ¿Qué ocurre con $i[A, B]$?

25. Supóngase que $i\hbar(\partial\psi/\partial t) = H\psi$. Demuéstrese que $\langle\psi, H\psi\rangle$ es constante en el tiempo. Este resultado se conoce como ***conservación de energía***.

27. Determínense las autofunciones en un pozo profundo si reemplazamos $[0, l]$ por $[-l, l]$.

28. Si A es simétrico, demuéstrese entonces que $\langle A\psi, \psi\rangle$ es real para cualquier $\psi \in V$. Interprétese este resultado usando el ejercicio 3, §**10.9**.

29. En la segunda demostración del teorema de convergencia puntual, usamos la igualdad

$$\int_0^\infty \frac{\operatorname{sen} x}{x}\, dx = \frac{\pi}{2}$$

(recuérdese que esta integral es condicionalmente convergente). Demuéstrese esta igualdad como sigue. Sea

$$F(t) = \int_0^\infty e^{-tx} \frac{\operatorname{sen} x}{x}\, dx.$$

Muéstrese que

$$F'(t) = -\int_0^\infty e^{-tx} \operatorname{sen} x\, dx = -(t^2 + 1)^{-1}$$

(véase el ejemplo resuelto **9.2** al final del capítulo 9) y por lo tanto, que $F(t) = -\tan^{-1} t + C$. Muéstrese que $F(t) \to 0$ cuando $t \to \infty$, de modo que $C = \pi/2$. Después analícese $F(0)$ para el resultado. La principal dificultad es la justificación de estos pasos (esta integral también se puede evaluar con métodos de variable compleja; véase J. Marsden y M. Hoffman, *Basic Complex Analysis*, 2ª ed., capítulo 4).

30. La derivada de la función delta se define como $\int_{-\infty}^{\infty} \delta'(x)f(x)dx = -f'(0)$ (véase §8.7). Calcúlese la transformada de Fourier de δ'; ¿cómo se relaciona con la de δ?

31. Las convoluciones se definen en el ejemplo resuelto **10.4**. Muéstrese que $f * g = g * f$. Úsese ese ejemplo para escribir la relación de Parseval para $f * g$.

32. **El teorema de Riesz-Fischer** Este teorema representa uno de los primeros y más importantes logros de la integral de Lebesgue. Para este problema no suponemos ningún conocimiento de la integral de Lebesgue, sino que damos por hecho que el conjunto de todas las funciones de cuadrado integrable es un espacio de Hilbert. Suponiendo esto, el teorema de Riesz-Fischer es bastante sencillo. A veces, según se vean las cosas, ¡el enunciado que acabamos de tomar como un hecho se llama teorema de Riesz-Fischer!

a. Demuéstrese el siguiente teorema.

Teorema de Riesz-Fischer *Sea V un espacio de Hilbert y $\varphi_0, \varphi_1, \ldots$ un conjunto ortonormal completo. Sean $c_0, c_1, \ldots$ números complejos y supóngase que $\sum_{n=0}^{\infty} |c_n|^2 < \infty$. Entonces existe $f \in V$ tal que*

$$\langle f, \varphi_i \rangle = c_i.$$

Así, cada serie

$$\sum_{n=0}^{\infty} c_n \varphi_n \quad con \quad \sum_{n=0}^{\infty} |c_n|^2 < \infty$$

es la serie de Fourier de alguna f.

[Sugerencia: sean $f_n = \sum_{i=0}^{n} c_i \, \varphi_i$ y muéstrese que f_n es una sucesión de Cauchy, mostrando que $\|f_m - f_n\|^2 = \sum_{i=n+1}^{m} |c_i|^2$.]

b. Úsese a para demostrar que para cada sucesión c_k, $k = 0, \pm 1, \pm 2, \ldots$, con

$$\sum_{-\infty}^{\infty} |c_k|^2 = \lim_{N \to \infty} \sum_{-N}^{N} |c_k|^2 < \infty,$$

existe una función f de cuadrado integrable en $[0, 2\pi]$ (o $[-\pi, \pi]$) tal que

$$f = \sum_{-\infty}^{\infty} c_k \varphi_k,$$

donde $\varphi_k = e^{ikx}$ y la convergencia es convergencia en media.

c. ¿Es

$$\sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} (1/n)e^{inx}$$

la serie de Fourier de alguna función integrable en $[-\pi, \pi]$? ¿Lo es

$$\sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} (1/\sqrt{n})e^{inx}?$$

33. Si $f\colon [a, b] \to \mathbb{R}$ tiene una discontinuidad solamente en $x_0 \in]a, b[$ y f' está acotada en $]x_0, b[$ y en $]a, x_0[$, demuéstrese entonces que f es una función de variación acotada. Muéstrese que se puede aplicar el teorema de Jordan-Dirichlet.

34. Sea $f(0) = 0, f(x) = x\operatorname{sen}(1/x)$ si $0 < x \leq \pi, f(x) = (2\pi - x)\operatorname{sen}(1/(2\pi - x))$ si $\pi \leq x < 2\pi$ y $f(2\pi) = 0$. Muéstrese que f es continua y periódica en $[0, 2\pi]$ y que la serie de Fourier de f converge en cada punto, pero que las hipótesis de los teoremas de Jordan y de Jordan-Dirichlet no se cumplen.

35. Investíguese la naturaleza de la convergencia de la serie de Fourier de cada una de las siguientes funciones en $]-\pi, \pi]$:

a. $f(x) = x^3$

b. $f(x) = (\operatorname{sen} x)^2$

c. $f(x) = \begin{cases} x^3 & x > 0 \\ -x^2 & x \leq 0 \end{cases}$

d. $f(x) = \begin{cases} x^2 + 1 & x \leq 0 \\ 0 & x < 0 \end{cases}$

e. $f(x) = \begin{cases} 1 & x \leq 0 \\ x^2 \operatorname{sen}\dfrac{1}{x} & x > 0 \end{cases}$

36. Supóngase que f es real y de cuadrado integrable en $[-l, l]$, y que

$$a_n = \frac{1}{l}\int_{-l}^{l} f(x)\cos\frac{n\pi x}{l}\,dx,\ \ n = 0, 1, 2, \ldots$$
$$b_n = \frac{1}{l}\int_{-l}^{l} f(x)\operatorname{sen}\frac{n\pi x}{l}\,dx,\ \ n = 1, 2, \ldots.$$

Muéstrese que

$$\frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2 + \sum_{n=1}^{\infty} b_n^2 = \frac{1}{l}\int_{-l}^{l} f(x)^2\,dx.$$

37. Fórmese la función

$$\varphi(x) = a\operatorname{sen} x + b\operatorname{sen}2x + c\operatorname{sen}3x$$

en $[0, \pi]$. ¿Para qué valores de a, b, c está φ más cerca (en media) de la función constante 1? ¿Qué ocurre en el intervalo $[-\pi, \pi]$?

38. Sea $g :]a, b[\to \mathbb{R}$ continua y supóngase que $g(a^+)$ y $g(b^-)$ existen. Demuéstrese que g está acotada. ¿Qué dice esto de la definición de una función continua a trozos?

39. **a.** Supóngase que f es derivable para $x > x_0$ y que $\lim_{x\to x_0^+} f'(x)$ existe. Muéstrese entonces que $f(x_0^+)$ también existe.

b. Si $\lim_{x\to x_0^+} f'(x)$ existe, muéstrese que es igual a

$$\lim_{h\to 0^+} \left\{\frac{f(x_0+h) - f(x_0^+)}{h}\right\}$$

c. Considérense las funciones $f_1(x) = x\operatorname{sen}(1/x)$, $f_2(x) = x^2\operatorname{sen}(1/x)$, $f_3(x) = x^3\operatorname{sen}(1/x)$ para $x > 0$. ¿Cuáles de los valores $\lim_{x\to 0^+} f_i'(x)$ y $f_i'(0^+)$ existen?

40. Sea x_0 una discontinuidad de salto para una función f, y sean $h(x) = f(x_0 - x)$ y $g(x) = f(x_0 + x)$. Muéstrese que $h(0^+) = f(x_0^-)$ y $g(0^+) = f(x_0^+)$.

41. Sea $f : [a, b] \to \mathbb{R}$ una función con una discontinuidad de salto en $x_0 \in]a, b[$, tal que $f'(x_0^+)$ y $f'(x_0^-)$ existen. Sea $g :]a, b[\to \mathbb{R}$ derivable y supóngase que

$$g(x) = f(x)\varphi(x).$$

Demuéstrese que g tiene una discontinuidad de salto en x_0, que $g'(x_0^+)$ y $g'(x_0^-)$ existen, y que

$$g(x_0^+) = f(x_0^+)\varphi(x_0)$$
$$g(x_0^-) = f(x_0^-)\varphi(x_0).$$

Aplíquese esto cuando

$$\varphi(x) = \frac{(x-x_0)/2}{\operatorname{sen}[(x-x_0)/2]}, \quad \text{si } x \neq x_0 \text{ y } \varphi(x_0) = 1,$$

para completar la demostración del teorema de la convergencia puntual.

42. Si $f_n \to f$ en media en $[a, b]$, demuéstrese entonces que $f_n \to f$ en media en cualquier subintervalo.

43. Supóngase que $f_n \to f$ y $g_n \to g$ en media en $[a, b]$. Defínase la función h como $h(x) = \int_a^x f(y)g(y)\,dy$ y sean h_n definidas de forma análoga como $h_n(x) = \int_a^x f_n(y)g_n(y)\,dy$. Demuéstrese que $h_n \to h$ uniformemente.

44. Establézcanse las fórmulas **5** y **6** de la tabla 10.5-4.

45. Establézcanse las siguientes fórmulas:

a. $x \operatorname{sen} x = 1 - (\cos x)/2 - 2\sum_1^\infty ((-1)^n \cos nx)/(n^2 - 1), \quad -\pi \leq x \leq \pi.$

b. $\log[\operatorname{sen}(x/2)] = -\log 2 - \sum_1^\infty (\cos nx)/n, \quad$ en $]0, 2\pi[$.

46. Aplíquese la relación de Parseval a las fórmulas **4a** y **6** de la tabla 10.5-4 para obtener algunas identidades aritméticas. ¿Qué justifica la interpretación de a_n y b_n como los coeficientes de $\cos nx$ y $\operatorname{sen} nx$?

47. Analícese el fenómeno de Gibbs para $f(x) = 2, x \geq 0; f(x) = 0, x < 0$.

48. Sea f con una discontinuidad de salto en x_0 tal que $f'(x_0^+)$, $f'(x_0^-)$ existen y supóngase que $f'(x)$ existe y es continua para $x \in]x_0 - \varepsilon, x_0[$ y $x \in]x_0, x_0 + \varepsilon]$. Muéstrese que f "exhibe un fenómeno de Gibbs" en x_0 y que la "sobreestimación" $\approx (f(x_0^+) - f(x_0^-))$. (1.179).

49. Úsese la serie de Fourier de $|\operatorname{sen} x|$ en la tabla 10.5-4 para mostrar que

$$\sum_1^\infty \frac{1}{4n^2 - 1} = \frac{1}{2} \quad \text{y} \quad \sum_1^\infty \frac{(-1)^n}{4n^2 - 1} = \frac{1}{2} - \frac{\pi}{4}.$$

50. Úsese la serie de senos de Fourier en $[0, \pi]$ para mostrar que

$$\cos \pi x = \frac{8}{\pi} \sum_{n=1}^\infty \frac{n}{4n^2 - 1} \operatorname{sen}(2\pi n x), \quad 0 < x < 1.$$

51. En la tabla 10.5-4, establézcanse las discontinuidades de todas las funciones (incluyendo los extremos de los intervalos) y los valores de la serie de Fourier en esos puntos.

52. Obténgase la fórmula **2a** de la tabla 10.5-4 observando que $f(x) = (x + |x|)/2$ y que la serie de $|x|$ es justamente la serie de cosenos en un semiintervalo para x.

53. Úsese el teorema de Jordan en $x = 0$ aplicado al ejemplo **10.5.5** para demostrar que $\pi^2/8 = \sum_{n=1}^\infty 1/(2n-1)^2$.

54. **a.** Sea f suave (C^∞) en $[-\pi, \pi]$ y supóngase que $f(-\pi) = f(\pi)$, $F^{(k)}(-\pi) = f^{(k)}(\pi)$, $k = 1, 2, \ldots$. Demuéstrese que la serie de Fourier de f se puede derivar cualquier número de veces y seguirá convergiendo uniformemente.

b. Muéstrese que para cualquier entero p, $n^p a_n \to 0$ y $n^p b_n \to 0$, donde a_n, b_n son los coeficientes de Fourier de f.

55. a. Sea $f : [0, \pi] \to \mathbb{R}$ continua y f' continua a trozos con discontinuidades de salto. Muéstrese entonces que la serie de cosenos en un semiintervalo de f converge uniforme y absolutamente a f.

b. Justifíquese la misma conclusión para la serie de senos en un semiintervalo si también suponemos que $f(0) = f(\pi) = 0$. Razónese la respuesta.

c. Muéstrese que sin la condición $f(-\pi) = f(\pi)$ del teorema **10.6.1**, la conclusión es falsa.

56. Sea

$$f(x) = \begin{cases} 0 & -\pi \leq x < 0 \\ 1 & 0 \leq x \leq \pi. \end{cases}$$

Muéstrese que si derivamos la serie de Fourier de f, obtenemos la serie de Fourier de $\delta(x) - \delta(x + \pi)$ (δ es la función delta de Dirac). ¿Puede el lector explicar en qué sentido $f' = \delta$?

57. Proporciónese el teorema que se obtiene al combinar el teorema **10.6.1** y el corolario **5.3.4,** y muéstrese por qué el teorema **10.6.2** es mejor.

58. Úsese el teorema **10.6.2** para obtener un teorema de derivación para la serie de cosenos en un semiintervalo.

59. Si f es de cuadrado integrable en $[-\pi, \pi]$ con coeficientes de Fourier a_n, b_n, demuéstrese entonces que $\sum_{n=1}^{\infty} a_n/n$ y $\sum_{n=1}^{\infty} b_n/n$ convergen absolutamente.

60. a. Sea $f : [-\pi, \pi] \to \mathbb{R}$ de cuadrado integrable y $g(x) = \int_{-\pi}^{x} f(y)\,dy$. Determínese el desarrollo de Fourier de g y establézcase el dominio donde es válido.

b. Repítase la parte **a** para la serie de cosenos en un semiintervalo de $f : [0, \pi] \to \mathbb{R}$.

61. En el ejemplo resuelto **10.3**, muéstrese que g es de variación acotada (por lo tanto, se puede aplicar el teorema de Jordan-Dirichlet).

62. Úsese el ejemplo resuelto **10.3** para determinar la serie de Fourier de x^3 en $[0, 2\pi]$ usando la de x^2 de la tabla 10.5-4.

63. Para cada una de las siguientes funciones, determínese el tipo de convergencia de la serie de Fourier y si la serie puede o no derivarse término a término.

a. $f(x) = x^3$ en $[-\pi, \pi]$

b. $$f(x) = \begin{cases} 3 & -\pi \leq x < -(1/2) \\ x & -(1/2) \leq x < 1 \\ 1 & 1 \leq x \leq \pi \end{cases}$$

c. $f(x) = \pi^3 - |x|^2$ en $[-\pi, \pi]$

d. $f(x) = \begin{cases} x^3 + 8 & -\pi \leq x \leq 0 \\ x^2 + 8 & 0 \leq x \leq \pi \end{cases}$

e. $f(x) = \begin{cases} 0 & -\pi \leq x \leq 0 \\ x^3 \operatorname{sen}(1/x) & 0 < x \leq \pi \end{cases}$

f. $f(x) = \begin{cases} -2n(x+1)x + 2n + 1 & 1/(n+1) \leq x \leq 1/n \\ 0 & -\pi \leq x \leq 0 \\ 1 & \text{resto de los valores de } [-\pi, \pi] \end{cases}$

g. $\log|\operatorname{sen}(x/2)|$ en $]-\pi, \pi[$

64. Supóngase que f es de cuadrado integrable en $[0, 2\pi]$ con coeficientes de Fourier a_n, b_n. Si $f(0^+)$, $f(0^-)$, $f'(0^+)$ y $f'(0^-)$ existen, demuéstrese que $\sum_{n=0}^{\infty} a_n$ converge. ¿Qué hipótesis garantizan que $\sum_{n=1}^{\infty} b_n$ converge?

Los ejercicios 65–71 se basan en **§10.7.**

65. Si la temperatura en los extremos de una barra se mantiene constante igual a cero, muéstrese que la temperatura T después de un tiempo t, si T es igual a f en $t = 0$, está dada por

$$T(x, t) = \sum_{n=1}^{\infty} b_n e^{-(n^2\pi^2 t/l)} \operatorname{sen} \frac{n\pi x}{l},$$

donde b_n son los coeficientes de la serie de senos en un semiintervalo para f.

66. Muéstrese que la solución de la ecuación del calor siempre es C^∞ para $t > 0$ (véase el teorema **10.7.4**).

67. **a.** En $[0, \pi] \times [0, \pi]$, encuéntrese una función φ tal que $\nabla^2\varphi = 0$ y

$$\varphi(x, 0) = \left(\frac{x - \pi}{2}\right)^2 - \frac{\pi^2}{4}, \ \varphi(0, y) = 0, \ \varphi(\pi, y) = 0, \ \varphi(x, \pi) = 0.$$

Explíquese la forma en que se alcanzan los valores en la frontera.

b. Repítase el problema con $\varphi(x, 0) = x$.

68. Muéstrese que la solución del texto es correcta para la ecuación de ondas bidimensional. Obténganse las soluciones fundamentales mediante separación de variables. El lector tal vez quiera utilizar el ejercicio 18.

69. En el teorema **10.7.2**, demuéstrese que

$$\lim_{t \to \infty} \sum_{n=1}^{\infty} a_n e^{-(n^2\pi^2 t/l)} \left(\cos \frac{n\pi x}{l}\right) = 0$$

(uniformemente en x).

70. En el teorema **10.7.4iii**, muéstrese que φ converge uniformemente en cualquier conjunto compacto en el interior del cuadrado.

71. **Problemas de contorno en ecuaciones diferenciales ordinarias**
Supóngase que f en $[-\pi, \pi]$ cumple $f(-\pi) = f(\pi)$ y que tiene los coeficientes de Fourier a_n, b_n. El hecho de que f' tenga los coeficientes $\alpha_n = nb_n$ y $\beta_n = na_n$ (véase el teorema **10.6.1**) es útil para resolver ciertos tipos de problemas de contorno.

a. Resuélvase la ecuación $f''(x) + kf(x) = g(x)$ para g dada en $[-\pi, \pi]$, si pedimos que $f'(-\pi) = f'(\pi)$ y que $f(-\pi) = f(\pi)$, observando que $-n^2a_n + ka_n = \tilde{a}_n$ y $-n^2b_n + kb_n = \tilde{b}_n$, donde $\tilde{a}_n$ y $\tilde{b}_n$ son los coeficientes de g. Por lo tanto, muéstrese que

$$f(x) = \frac{\tilde{a}_0}{2k} + \sum_{n=1}^{\infty} \frac{\tilde{a}_n}{k - n^2} \cos nx + \frac{\tilde{b}_n}{k - n^2} \operatorname{sen} nx.$$

b. Resuélvase la ecuación del apartado anterior para $[-l, l]$.

72. Sea δ un número real dado, $0 < \delta < \pi$. Defínase

$$f(x) = \begin{cases} 1 & |x| \le \delta \\ 0 & \delta < |x| \le \pi. \end{cases}$$

a. Calcúlese la serie de Fourier de f.

b. Evalúese en $x = \pi$ y muéstrese que

$$\frac{\delta}{2} = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{\operatorname{sen} k\delta}{k}.$$

c. ¿Qué dice la relación de Parseval en el caso de f?

73. Evalúese

$$\lim_{k \to \infty} \int_0^{\pi} \sqrt{x} \operatorname{sen}^2 kx \, dx.$$

74. Verifíquese que el producto infinito del ejemplo resuelto **10.5**,

$$f(\lambda) = \lambda \prod_{n=1}^{\infty} \left(1 - \frac{\lambda^2}{n^2}\right),$$

es periódico de periodo 2, $f(\lambda + 2) = f(\lambda)$. Muéstrese primero que $f(\lambda + 1) = -f(\lambda)$.

75. (W.A.J. Luxemburg) Sea φ_n un conjunto de funciones ortonormales en L^2 del intervalo $[a, b]$. Muéstrese que

a. φ_n es completo sii $x - a = \sum_{n=1}^{\infty} \left| \int_a^x \varphi_n(t)\, dt \right|^2$ para todo $x \in [a, b]$ y que

b. φ_n es completo sii

$$\frac{(b-a)^2}{2} = \sum_{n=1}^{\infty} \int_a^b \left| \int_a^x \varphi_n(t)\, dt \right|^2 dx.$$

76. (Demostración de Lebesgue de la completitud del sistema trigonométrico, extraída de A. Zygmund, *Trigonometric Series*, Chelsea, Nueva York) Proporciónese otra demostración del teorema de completitud media como sigue. Sea $f : [-\pi, \pi] \to \mathbb{R}$ continua y ortogonal a $\cos nx$, $\operatorname{sen} mx$. Demuéstrese que $f = 0$ como sigue. Supóngase que $f(x) > \varepsilon$ para $x \in I =]x_0 - \delta, x_0 + \delta[$. Sea $T_n(x) = [t(x)]^n$, donde $t(x) = 1 + \cos(x - x_0) - \cos \delta$, y muéstrese que $T_n(x) \geq 0$ en I, que $T_n(x) \to \infty$ uniformemente en cada subintervalo cerrado de I y que los T_n están uniformemente acotados fuera de I. Úsese esto para mostrar que $\langle f, T_n \rangle = 0$ es imposible para n grande. Para f arbitraria, trátese de aplicar la relación recién obtenida a $F(x) = \int_{-\pi}^{x} f(t)dt$.

Apéndice A

Ejercicios diversos

Este apéndice contiene ejercicios complementarios. Los ejercicios 1–53 son ejercicios diversos que analizan una variedad de temas. La mayoría no son preguntas de rutina, sino que piden ejemplos, determinación de la certeza o falsedad de ciertas afirmaciones, o desarrollo de variantes o extensiones de ideas del texto. Los ejercicios 54–66 son preguntas típicas de examen que pueden ser útiles como repaso.

Ejercicios

1. Para dos conjuntos $A, B \subset \mathbb{R}^n$, defínase la distancia entre ellos como $d(A, B) = \inf\{d(x, y) \mid x \in A, y \in B\}$.

a. Encuéntrense dos conjuntos cerrados tales que $A \cap B = \varnothing$ y, sin embargo, $d(A, B) = 0$.

b. Para A compacto y $x \notin A$, muéstrese que $d(A, x) = d(y, x) > 0$ para algún $y \in A$.

c. Muéstrese que para cada conjunto $A \subset \mathbb{R}^n$ tenemos $\text{cl}(A) = \{x \in \mathbb{R}^n \mid d(x, A) = 0\}$.

2. Sea $A \subset \mathbb{R}^n$. Muéstrese que A es compacto sii toda función continua $f : A \to \mathbb{R}$ está acotada superiormente y alcanza su máximo en algún punto en A.

3. Muéstrese que $f : \mathbb{R} \to \mathbb{R}$ tiene una derivada continua sii el límite doble

$$\lim_{(x,y)\to(x_0,x_0)} \frac{f(x) - f(y)}{x - y}$$

existe para cada $x_0 \in \mathbb{R}$.

4. Muéstrese que $\sum_{n=1}^{\infty}(x+1)^2/n^2$ y $\sum_{n=1}^{\infty}(x+1)^n/n!$ convergen ambas uniformemente en $[0, 1]$.

5. ¿Qué es incorrecto en el siguiente argumento?

$$f(x) = 1/(1+x)^2 = (d/dx)[-1/(1+x)],$$

de modo que

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{a\to\infty} \int_{-a}^{a} f(x)\,dx = \lim_{a\to\infty}\left[\frac{-1}{1+a} - \frac{-1}{1-a}\right] = 0.$$

6. Proporciónese un ejemplo de una función continua $f: \mathbb{R} \to \mathbb{R}$ tal que $f(\mathbb{R})$ no sea cerrado. Si A es un intervalo cerrado y acotado en $\mathbb{R}$, ¿debe $f(A)$ ser cerrado?

7. En primera instancia, se podría pensar que los intervalos $]0, 1[$ y $[0, 1]$ son muy similares. Enúnciense al menos cinco diferencias significativas entre ellos en términos de la topología y las funciones continuas.

8. **a.** Encuéntrese un ejemplo de un conjunto cerrado A en $\mathbb{R}^n$ tal que $A = \partial(A)$.

b. Si $A = \partial(A)$, muéstrese que A es cerrado.

c. Demuéstrese que si $\text{int}(A) = \varnothing$, entonces ($A = \partial(A) \Leftrightarrow A$ es cerrado).

d. Demuéstrese que si A es cerrado y $A \subset \partial(A)$, entonces $A = \partial(A)$ e $\text{int}(A) = \varnothing$.

e. Encuéntrese un conjunto A tal que $A \subset \partial(A)$ pero $A \neq \partial(A)$.

9. ¿Son ciertos o falsos los siguientes enunciados? (Todos los conjuntos tienen volumen y todas las funciones están acotadas y son integrables.)

a. Si $A \supset B$ y $A\backslash B$ tiene medida nula, entonces $\int_A f = \int_B f$.

b. Si $\{x \mid f(x) \neq g(x)\}$ tiene medida nula, entonces $\int_A f = \int_B f$.

c. Si $f \geq 0$, $g \geq 0$ y $\int_A f = \int_A g$, entonces $f = g$ en A (excepto posiblemente en un conjunto de medida nula).

d. La misma pregunta que en **c**, con la hipótesis adicional de que $f \geq g$.

10. Sea $f: [a, b] \to \mathbb{R}$ integrable.

a. Demuéstrese que $F(x) = \int_a^x f(t)\,dt$ es uniformemente continua en $[a, b]$.

b. Muéstrese que F tiene derivada en x_0 si f es continua en x_0.

c. Muéstrese que F es derivable excepto en un conjunto de medida nula.

11. **a.** Sea $T: \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal. Demuéstrese que T preserva la norma (es decir, $\|Tx\| = \|x\|$) sii T preserva el producto interno, $\langle Tx, Ty\rangle = \langle x, y\rangle$ (véase el ejercicio 12, capítulo 1).

b. Si T preserva la norma (o el producto interno), entonces T es un isomorfismo.

12. Demuéstrese el principio de Cavalieri: si A, $B \subset \mathbb{R}^3$ tienen volumen, y si cada plano paralelo al plano XY interseca A y B dando lugar a conjuntos con áreas iguales, entonces A y B tienen el mismo volumen. [Sugerencia: úsese el teorema de Fubini.]

Observación: en relación con este problema, véase Gelbaum y Olmsted, *Counterexamples in Analysis*, ejemplo 6, capítulo 11. Para las aplicaciones, véase un texto de cálculo.

13. Muéstrese que un conjunto $A \subset \mathbb{R}^n$ tiene volumen sii para cada $\varepsilon > 0$ existen un conjunto $V_\varepsilon \subset A$ y un conjunto $W_\varepsilon \supset A$ tales que V_ε y W_ε tienen volumen y $v(W_\varepsilon \setminus V_\varepsilon) = v(W_\varepsilon) - v(V_\varepsilon) < \varepsilon$. Muéstrese que si la última condición se cumple, entonces el volumen de A es $\inf\{v(W_\varepsilon) \mid \varepsilon > 0\} = \sup\{v(V_\varepsilon) \mid \varepsilon > 0\}$.

14. Muéstrese que el volumen de la figura obtenida al rotar el área por debajo de la gráfica de una función no negativa $f: [a, b] \to \mathbb{R}$ una vez en torno del eje X está dado por $\int_a^b \pi f(x)^2\, dx$. Úsese esta fórmula para calcular el volumen de $\{(x, y, z) \in \mathbb{R}^3 \mid 1 < x < 2 \text{ e } y^2 + z^2 < x^4$.

15. Sea $f: [a, b] \to \mathbb{R}$ continua y supóngase que $f(a)f(b) < 0$. Muéstrese que existe $x \in]a, b[$ tal que $f(x) = 0$.

16. Sea $f: \mathbb{R}^n \to \mathbb{R}^n$ de clase C^1 y supóngase que $Jf(x) \neq 0$ para todo x. Sean $x_0 \in \mathbb{R}^n$ y $B = \{x \in \mathbb{R}^n \mid f(x) = x_0\}$. Muéstrese que B no tiene puntos de acumulación.

17. Proporciónese un ejemplo de una función f de clase C^1 con derivada igual a cero en un punto x pero que sea inyectiva en una vecindad de x. Muéstrese que f^{-1} no puede ser derivable en $f(x)$.

18. Muéstrese que si los conjuntos A y B tienen volumen, también $A \cup B$. Si $A \cap B$ y $A \setminus B$ también tienen volumen, demuéstrese que:

a. $v(A \cup B) = v(A) + v(B) - v(A \cap B)$.

b. $v(A \setminus B) = v(A) - v(B)$ si $A \supset B$.

19. Supóngase que $f: A \subset \mathbb{R}^n \to \mathbb{R}$ es integrable y que $f = g$ excepto en un conjunto de medida nula. Muéstrese que g es integrable. Muéstrese que esto es falso si reemplazamos "contenido nulo" por "medida nula".

20. **a.** Muéstrese que si $\{U_\alpha\}$ es una familia de conjuntos abiertos disjuntos en $\mathbb{R}^n$, entonces la familia es numerable. [Sugerencia: elíjase un punto de coordenadas racionales en cada conjunto y úsese el hecho de que el conjunto de tales puntos es numerable.]

b. Si A es un subconjunto abierto de $\mathbb{R}^n$, muéstrese que las componentes conexas de A son abiertas y que hay una cantidad numerable de ellas.

c. Demuéstrese que cada conjunto abierto en $\mathbb{R}$ es la unión de una colección numerable de intervalos.

21. Sea $f: [a, b] \to [\alpha, \beta]$ una función estrictamente creciente en $[a, b]$ y suprayectiva $[\alpha, \beta]$. Muéstrese que f y f^{-1} son continuas.

22. Demuéstrese el lema del recubrimiento de Lebesgue: sea A un subconjunto compacto de $\mathbb{R}^n$ y sea $\{V_\alpha\}$ un recubrimiento abierto de A; entonces, existe $\varepsilon > 0$ tal que si S es cualquier rectángulo contenido en A con lados menores que ε, entonces $S \subset V_{\alpha_0}$ para algún α_0.

23. Sea $f: A \subset \mathbb{R}^n \to \mathbb{R}$ y sea $M \subset A$ el conjunto de máximos locales (estrictos) de f. Muéstrese que $f(M)$ es finito o numerable. Proporciónense ejemplos.

24. Sea S un conjunto abierto conexo en $\mathbb{R}^n$. Sea A una componente conexa de $\mathbb{R}^n \backslash S$. Muéstrese que $\mathbb{R}^n \backslash A$ es conexo.

25. Encuéntrese una función continua no constante $f: [a, b] \to \mathbb{R}$ que tenga su máximo en $x_0 \in]0, 1[$ pero para la cual no exista $f'(x_0)$.

26. Un conjunto $B \subset A$ es denso en A si $A \subset \text{cl}(B)$. Muéstrese que esto es equivalente a la condición de que para cada conjunto abierto U tal que $A \cap U \neq \varnothing$ se tiene $B \cap U \neq \varnothing$. ¿Es A denso en $\text{cl}(A)$? Muéstrese que $\mathbb{R}^n$ tiene un subconjunto denso numerable.

27. a. Sea $A \subset \mathbb{R}^n$ un conjunto con volumen. Muéstrese que $\text{int}(A)$ y $\text{cl}(A)$ tienen volumen y que $v(A) = v(\text{int}(A)) = v(\text{cl}(A))$.

b. Muéstrese que si A es un conjunto tal que $\text{cl}(A)$ tiene volumen, no podemos concluir que el propio A tenga volumen.

c. Muéstrese que si A es un conjunto tal que $\text{int}(A) = \varnothing$, no podemos concluir que A tenga contenido nulo o medida nula.

28. Sea $g: A \subset \mathbb{R}^n \to B \subset \mathbb{R}^n$ de clase C^1 en un conjunto abierto A y supóngase que $B = g(A)$. Decimos que g preserva el volumen si para cada conjunto $D \subset A$ tal que D y $g(D)$ tienen volumen, tenemos $v(g(D)) = v(D)$. Supóngase que g es inyectiva y que $Jg(X) \neq 0$ en cada $x \in A$; muéstrese entonces que g preserva el volumen sii $|Jg(x)| = 1$ para todo $x \in A$.

29. a. Muéstrese que si un conjunto A tiene contenido nulo, entonces $\text{cl}(A)$ tiene contenido nulo. ¿Es cierto esto para medida nula?

b. Sea $f: A \subset \mathbb{R}^n \to \mathbb{R}^m$ de clase C^1 en un conjunto abierto A. Sea $B \subset A$ y $\text{cl}(B) \subset A$ y supóngase que B es compacto. Si B tiene contenido nulo o medida nula, muéstrese que también ocurre lo mismo para $f(B)$. [Sugerencia: considérese el caso en que $Jf(x) = 0$ por separado y úsese el teorema de Sard (ejercicio 27, capítulo 9).]

30. Un subconjunto A de $\mathbb{R}^n$ es homeomorfo a un subconjunto B de $\mathbb{R}^m$ si existe una transformación continua $\varphi : A \to B$ con inversa continua $\varphi^{-1} : B \to A$. Decimos que φ es un homeomorfismo.

a. Encuéntrese un ejemplo de una biyección $\varphi : A \to B$ que sea continua pero que no sea un homeomorfismo.

b. Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ continua y $\Gamma = \{x, f(x) \in \mathbb{R}^n \times \mathbb{R}^m \mid x \in A\}$ la gráfica de f. Muéstrese que A y Γ son homeomorfos.

31. Sea $f : [a, b] \to \mathbb{R}$ una función monótona (digamos, por ejemplo, que f no es decreciente: $x \le y \Rightarrow f(x) \le f(y)$).

a. Muéstrese que los límites izquierdo y derecho, $f(x^+) = \lim_{h\to 0^+} f(x + h)$ y $f(x^-) = \lim_{h\to 0^+} f(x - h)$, existen en cada x de $[a, b]$ (sólo uno de los dos en los extremos).

b. Muéstrese que f tiene como mucho una cantidad numerable de discontinuidades. [Sugerencia: sea P_n el conjunto de puntos en los que el salto de f excede a $1/n$; muéstrese que P_n es finito y considérese la unión de todos los P_n para $n = 1, 2, 3, \ldots$.]

c. Un famoso teorema de Lebesgue afirma que para tal f, la derivada de f existe, excepto quizás en un conjunto de medida nula. Considérense algunos ejemplos para verificar la validez de los resultados. Revísese una demostración en, por ejemplo, Hewitt y Stromberg, *Real and Abstract Analysis* y redáctese un breve escrito acerca de las características principales de la demostración.

32. Demuéstrese que la transformación

$$\left[\begin{array}{rl} x_1 = & u_1 \\ x_2 = & u_1 + u_2 \\ x_3 = & u_1 + u_2 + u_3 \\ & \vdots \\ x_n = & u_1 + \cdots + u_n \end{array}\right.$$

deja invariantes los volúmenes.

33. Sea $g : [0, 1] \to \mathbb{R}$ integrable. Muéstrese que

$$\int_0^1 \left[\int_x^1 g(t)\,dt\right] dx = \int_0^1 t g(t)\,dt.$$

34. Sea K un conjunto compacto. Si $\langle f_n \rangle_1^\infty$ es una sucesión uniformemente convergente de funciones continuas de valores reales en K, muéstrese que f_n es equicontinua. El recíproco no es cierto. Proporciónese un contraejemplo.

35. Sea S un subconjunto de $\mathbb{R}^n$ con volumen y sea $t > 0$. Sea R el conjunto de puntos $R = \{tx_1, \ldots, tx_n) \mid (x_1, \ldots, x_n) \in S\}$. Muéstrese que $v(R) = t^n v(S)$. ¿Qué ocurre si $t < 0$?

36. Explíquese cómo es posible que exista el fenómeno de Gibbs y que la serie de Fourier siga convergiendo en media y puntualmente.

37. **a.** Sea $f(x)$ una función con serie de Fourier $(a_0/2) + \sum_{n=1}^{\infty}[a_n \cos nx + b_n \operatorname{sen} nx]$ en $[-\pi, \pi]$. Defínase la reflexión de f como $g(x) = f(-x)$. Muéstrese que la serie de Fourier de g es $(a_0/2) + \sum_{n=1}^{\infty}[a_n \cos nx - b_n \operatorname{sen} nx]$.

b. Sea

$$f(x) = \begin{cases} 0, & \text{para } -\pi \le x \le 0, \\ x, & \text{para } 0 < x \le \pi. \end{cases}$$

Recuérdese que la serie de Fourier de f es

$$\frac{\pi}{4} + \sum_{n=1}^{\infty} \left(\frac{(-1)^n - 1}{\pi n^2} \cos nx - \frac{(-1)^n}{n} \operatorname{sen} nx \right).$$

Úsese el apartado **a** para mostrar que la serie de Fourier de $|x|$ en $[-\pi, \pi]$ es

$$\frac{\pi}{2} + 2\sum_{n=1}^{\infty} \frac{(-1)^n - 1}{\pi n^2} \cos nx = \frac{\pi}{2} - 4\sum_{k=1}^{\infty} \frac{\cos[(2k-1)x]}{\pi(2k-1)^2}.$$

c. Úsese el apartado **b** para mostrar que $\dfrac{\pi^2}{8} = 1 + \dfrac{1}{3^2} + \dfrac{1}{5^2} + \dfrac{1}{7^2} + \cdots$.

d. Úsese la parte **b** para obtener la serie de cosenos de Fourier de x en $[0, \pi]$ y viceversa.

38. **a.** Sean V un espacio con producto interno y $\varphi_0, \varphi_1, \ldots$ una sucesión ortonormal completa. Supóngase que W es un subespacio de V y que $f \in W$ sii $\langle f, \varphi_0 \rangle = 0$. Demuéstrese que $\varphi_1, \varphi_2, \ldots$ es un sistema ortonormal completo para W. Generalícese.

b. Aplíquese el apartado **a** al sistema trigonométrico y

i. $W = \{f : [-\pi, \pi] \to \mathbb{R} \mid \int_{-\pi}^{\pi} f(x)\,dx = 0\}$.
ii. $W = \{f : [-\pi, \pi] \to \mathbb{R} \mid f \text{ es par}\}$.
iii. $W = \{f : [-\pi, \pi] \to \mathbb{R} \mid f \text{ es impar}\}$.

39. Si $f : [a, b] \to \mathbb{R}$ es de cuadrado integrable, demuéstrese que f es integrable; es decir: $\int_a^b |f(x)|^2\,dx < \infty \Rightarrow \int_a^b |f(x)|\,dx < \infty$. [Sugerencia: úsese la desigualdad de Cauchy-Schwarz.]

40. **a.** Supóngase que $f : [-\pi, \pi] \to \mathbb{R}$ es continua a trozos y que sólo tiene discontinuidades de salto. Muéstrese que la suma de la serie de Fourier de f en x sólo depende de los valores de f en cualquier vecindad de x. Esto se denomina la *propiedad de localización de Riemann*. [Sugerencia: aplíquese el teorema **10.3.2.**]

b. Los coeficientes de Fourier de f dependen de los valores de f en todo $[-\pi, \pi]$. ¿Cómo se puede conciliar esto con el resultado del apartado **a**? [Sugerencia: estúdiese la demostración de **10.3.2.**]

41. Supóngase que la serie de Fourier de $f: [-\pi, \pi] \times [-\pi, \pi] \to \mathbb{R}$ es

$$\sum_{n,m=-\infty}^{\infty} c_{n,m} e^{inx} e^{imy}.$$

(Véase el ejercicio 18, capítulo 10.)

a. Desarróllese la serie de Fourier de f en forma trigonométrica.

b. Para y fijo, sea $g(x) = f(x, y)$. Muéstrese que los coeficientes de la serie exponencial de Fourier de g son

$$\hat{G}(n) = c_n = \sum_{m=-\infty}^{\infty} c_{n,m} e^{imy}.$$

c. Si f es de cuadrado integrable sabemos que su serie de Fourier converge a f en media (véase el teorema **10.3.1**). El propósito aquí es dar un teorema de convergencia puntual. Muéstrese que si f es de clase C^1 y $f(x, \pi) = f(x, -\pi)$ y $f(\pi, y) = f(-\pi, y)$ para todo x e y, entonces la serie de Fourier de f converge a $f(x, y)$ puntualmente. [Sugerencia: úsese el apartado **b** y **10.3.2.**]

42. ¿Para qué valores de p es $\sum_{n=1}^{\infty} (\operatorname{sen} nx)/n^p$ la serie de Fourier de una función de cuadrado integrable? (Véase el ejercicio 32, capítulo 10.)

43. **a.** Muéstrese que

$$\left| \sum_{k=0}^{n} \cos kx \right| \leq \left| \frac{1}{\operatorname{sen}(x/2)} \right|$$

para $x \neq 0$.

b. Considérese la serie de Fourier de la función escalonada

$$f(x) = \sum_{n=1}^{\infty} \frac{\cos nx}{n}.$$

Muéstrese que para cada $\delta > 0$, esta serie converge uniformemente en $[\delta, \pi]$. [Sugerencia: úsese el apartado **a** y el criterio de Dirichlet.]

c. Generalícese **b** a cualquier serie de Fourier $f(x) = \sum_{n=1}^{\infty} a_n \cos nx$ con a_n decreciente. Conclúyase que f debe ser continua en $]0, \pi[$.

d. Dedúzcase de **c** que si f tiene una discontinuidad en x_0 y $0 < x_0 < \pi$, entonces los coeficientes de Fourier de f no son decrecientes.

44. Una cuerda en $[0, l]$ se desplaza inicialmente en $t = 0$ mediante $f(x) = (x - l/2)^2 - l^2/4$. Determínese una fórmula para el desplazamiento tras un tiempo t.

45. a. Si una barra con extremos aislados tiene temperatura T constante en $t = 0$, muéstrese que T es constante para todo $t > 0$.

b. Si una barra en $[0, \pi]$ tiene temperatura en $t = 0$ dada por $\operatorname{sen} x$, encuéntrese la temperatura para $t > 0$.

c. Lo mismo que en **b,** excepto que $T = \cos x$ en $t = 0$.

46. a. Encuéntrese una función φ en $[0, \pi] \times [0, \pi]$ tal que $\nabla^2\varphi = 0$ y que $\varphi(x, 0) = \cos x$ y $\varphi(x, \pi) = 0 = \varphi(0, y) = \varphi(\pi, y)$.

b. En el apartado **a**, reemplácese $\varphi(0, y) = 0$ por $\varphi(0, y) = 1$ y determínese la función.

c. ¿En qué sentido se alcanzan los valores en la frontera en **a** y **b**?

47. Sea V un espacio con producto interno. Por lo general, $\|f_n\| \to \|f\|$ no implica $f_n \to f$ (ejercicio 16, capítulo 3). Sin embargo, muéstrese que $\|f_n\| \to \|f\|$ sí que implica $f_n \to f$ en media si f_n es la n-ésima suma parcial de la serie de Fourier de f con respecto de una familia ortonormal.

48. Sea f: $\mathbb{R} \to \mathbb{R}$ dos veces derivable. Muéstrese que

a. Si $F(x, y) = f(xy)$, entonces $x\, \partial F/\partial x = y \partial F/\partial y$.

b. Si $F(x, y) = f(ax + by)$, entonces $b\, \partial F/\partial x = a \partial F/\partial y$.

c. Si $F(x, y) = f(x^2 + y^2)$, entonces $y\, \partial F/\partial x = x \partial F/\partial y$.

d. Si $F(x, y) = f(x + cy) + f(x - cy)$, entonces $c^2\partial^2 F/\partial x^2 = \partial^2 F/\partial y^2$.

49. Demuéstrese el lema de Kronecker: si $\sum_{n=1}^{\infty}(x_n/n)$ converge, entonces $(x_1 + \cdots + x_n)/n \to 0$ cuando $n \to \infty$; (es decir, $x_n \to 0$ en el sentido de Cesaro).

50. Sea f: $[a, b] \to \mathbb{R}$ una función acotada integrable con $f(x) \geq m > 0$ para todo x en $[a, b]$. Muéstrese que $\left(\int_a^b (1/f)\right)\left(\int_a^b f\right) \geq (b - a)^2$.

51. Supóngase que f: $[0, 1] \to \mathbb{R}$ es integrable, que $\int_0^1 f(x)\, dx \geq 7$ y que $0 \leq f(x) \leq 10$ para todo x en $[0, 1]$. Defínase el conjunto $E = \{x \in [0, 1] \mid f(x) \geq 1\}$ y supóngase que E tiene volumen. Muéstrese que $v(E) \geq 1/2$.

52. Supóngase que $f : [0, 2\pi] \to \mathbb{R}$ es continua y que $f(0) = f(2\pi)$. Sea

$$s_N = \sum_{k=-N}^{N} \langle f, \varphi_k \rangle \varphi_k$$

la N-ésima suma parcial de la serie de Fourier de f y defínase

$$\Phi(y) = \int_0^y f(x)\,dx \qquad \text{y} \qquad \Sigma_N(y) = \int_0^y s_N(x)\,dx.$$

Establézcanse cuáles de las siguientes afirmaciones "deben ser ciertas" o "podrían ser falsas".

a. $\int_0^{2\pi} s_N(x)e^{ix}\,dx \to \int_0^{2\pi} f(x)\,dx$ cuando $N \to \infty$.

b. $\int_0^{2\pi} x^2 s_N(x)\,dx \to \int_0^{2\pi} x^2 f(x)\,dx$.

c. $\| s_N - f \| \to 0$.

d. $s_N(2) \to f(2)$.

e. Σ_N es la N-ésima suma parcial de la serie de Fourier de Φ.

f. $\| \Sigma_N - \Phi \| \to 0$.

g. $\Sigma_N(2) \to \Phi(2)$.

h. $\Sigma_N \to \Phi$ uniformemente en $[0, 2\pi]$.

53. *El núcleo de Poisson y las funciones armónicas.* Sea $f(\theta)$ continua y 2π-periódica (podemos imaginar f como una función definida sobre la circunferencia del círculo unidad en el plano). Por el teorema de Fejér, sabemos que la serie de Fourier de f converge a f en sentido $(C, 1)$. Dedúzcase que

$$\lim_{r \to 1^+} \sum_{k=-\infty}^{\infty} c_k r^{|k|} e^{ik\theta} = f(\theta).$$

(Obsérvese el exponente $|k|$ para los índices negativos.) Defínase

$$u(r, \theta) = \sum_{k=-\infty}^{\infty} c_k r^{|k|} e^{ik\theta}$$

para $0 \le r < 1$. Consideramos u como una función en el interior del disco unidad del plano. En coordenadas rectangulares x, y tenemos

$$u(x, y) = c_0 + \sum_{k=1}^{\infty} (c_k(x + iy)^k + c_{-k}(x - iy)^k).$$

Demuéstrese que esta serie converge *uniformemente* en cada disco de radio menor que 1.

Muéstrese (por la teoría general de las series de potencias) que podemos derivar u término a término cualquier número de veces. De esta forma demostramos que

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0;$$

es decir, u es una solución de la ecuación de Laplace (una *función armónica).* Ya hemos visto que $u(r, \theta) \to f(\theta)$ cuando $r \to 1^-$, de modo que hemos resuelto el "problema de Dirichlet": encontrar una función armónica en el disco unidad que tenga una función dada como condición de contorno.

Para $0 \leq r < 1$, demuéstrese que

$$u(r, \theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) P_r(\theta - t)\, dt$$

donde $P_r(\theta - t) = \sum_{-\infty}^{\infty} r^{|k|} e^{ik(\theta - k)}$. La función $P_r(y) = \sum_{-\infty}^{\infty} r^{|k|} e^{iky}$ es el *núcleo de Poisson*. Súmese esta serie explícitamente para demostrar que

$$P_r(y) = \frac{1 - r^2}{1 + r^2 - 2r \cos y}.$$

Muéstrese que este núcleo tiene las mismas propiedades cruciales que el núcleo de Fejér.

a. 2π-periodicidad.

b. $(1/2\pi) \int_{-\pi}^{\pi} P_r(t)\, dt = 1$.

c. $P_r(t) \geq 0$.

d. Para cada $\delta > 0$ fijo, tenemos

$$\lim_{r \to 1^-} \int_{\delta \leq |t| \leq \pi} P_r(t)\, dt = 0.$$

Dedúzcase que la función $u(r, \theta)$ analizada arriba converge a $f(\theta)$ *uniformemente* cuando $r \to 1^-$.

Preguntas tipo examen

Los ejercicios 54–58 son preguntas tipo examen, basadas en los capítulos 1–6 y 8.

54. a. Defínase el supremo de un conjunto S.

b. Encuéntrese $\sup\{x \in \mathbb{R} \mid x^2 + x < 3\}$.

c. ¿Qué se entiende al decir que $\mathbb{R}$ es completo?

d. Sea $\langle x_n \rangle_1^\infty$ una sucesión convergente en $\mathbb{R}$. Demuéstrese que es una sucesión de Cauchy.

e. Sea $0 < \alpha < 1$ y defínase $x_0 = \alpha$; inductivamente, defínase $x_n = (x_{n-1} + 1)/2$ para $n = 1, 2, 3, \ldots$. Demuéstrese que x_n converge a 1 cuando $n \to \infty$.

55. a. Defínase el término "conjunto conexo".

b. Defínase el término "conexo por arcos".

c. Establézcase y demuéstrese una versión general del teorema de los valores intermedios.

d. Demuéstrese que $\{(x, y) \in \mathbb{R}^2 \mid x = 0 \text{ e } y \geq 0\} \cup \{(x, y) \in \mathbb{R}^2 \mid x = y \text{ y } x > 0\}$ es conexo.

e. Si A y B son conjuntos conexos en $\mathbb{R}^n$ y $A \cap B \neq \emptyset$; muéstrese que $A \cup B$ es conexo.

56. a. Defínase lo que quiere decir que una sucesión de funciones $f_k : \mathbb{R}^n \to \mathbb{R}$ converja uniformemente.

b. Demuéstrese que $\sum_{k=1}^{\infty} (\text{sen } kx)^2 / k^{3/2}$ converge uniformemente para $x \in \mathbb{R}$.

c. ¿Es $f(x) = \sum_{k=1}^{\infty} (\text{sen } kx)^2 / k^{3/2}$ una función continua de x? Justifíquese la respuesta.

d. Sean $f_k(x) = (1/kx) + 1$ para $k = 1, 2, 3, \ldots$ y $x \in]0, 1[$. Demuéstrese que $f_k \to 1$ puntualmente.

e. ¿Converge uniformemente la sucesión $\langle f_k \rangle_1^\infty$ del apartado **d** en $]0, 1[$?

57. a. Sean $f_n : [a, b] \to \mathbb{R}$ funciones continuas y derivables en $]a, b[$ con f_n' continuas. Supóngase que f_n converge uniformemente a f y que f_n' converge uniformemente a g. Enúnciese un teorema relativo a la derivabilidad de f.

b. Demuéstrese el teorema enunciado en **a.** Enúnciense claramente todos los resultados utilizados.

c. Sean $f_k(x) = (\text{sen } kx)/k^2$ ¿Se aplica el teorema?

d. Enúnciese un resultado que garantice que la siguiente operación sea válida:

$$\sum_{k=1}^{\infty} \left(\int_a^b g_k(x)\, dx \right) = \int_a^b \left(\sum_{k=1}^{\infty} g_k(x) \right) dx.$$

e. Defínase $\sum_{n=0}^{\infty} x^n / n!$. Úsese el apartado **d** para demostrar que $\int_0^x e^t\, dt = e^x - 1$.

58. **a.** Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^n$, donde A es abierto. Proporciónese una definición de la derivada de f.

b. Para $f : A \subset \mathbb{R}^n \to \mathbb{R}$, defínase el gradiente de f y analícese el significado geométrico de $\langle \operatorname{grad} f(x), e \rangle$.

c. Si S es una superficie definida por $S = \{x \mid F(x) = c\}$ para alguna constante c, justifíquese el hecho de que grad $f(x)$ es perpendicular a S si $x \in S$.

d. Encuéntrese la ecuación del plano tangente a la superficie $x^2 + y^3 + z^4 = 3$ en $(1, 1, 1)$.

e. Justifíquese que las dos superficies $x^2 + y^2 + z^2 = 3$ y $x^3 + y^3 + z^4 = 3$ sean tangentes en el punto $(1, 1, 1)$.

Los ejercicios 59–62 se basan en los capítulos 6–10.

59. **a.** Sea $f : \mathbb{R}^n \to \mathbb{R}^m$. Defínase lo que quiere decir que f sea diferenciable en $x \in \mathbb{R}^n$.

b. ¿Es cierto que la existencia de las derivadas parciales implica que f es diferenciable? Analícese.

c. Sea $f : \mathbb{R}^2 \to \mathbb{R}^3, f(x, y) = (xy, e^y, \cos x)$. Calcúlese $Df(1, 0)$.

d. Sea $h(x, y) = f(g(x, y), k(y), p(x))$. Escríbase una fórmula para $\partial h/\partial x$. Justifíquese en términos de la regla de la cadena.

e. Sea $f :\ \mathbb{R} \to \mathbb{R}$ derivable. Supóngase que f y f' no tienen raíces comunes. Demuéstrese que f sólo tiene un número finito de ceros en $[0, 1]$.

60. **a.** ¿Qué dice el teorema de la función inversa para funciones $f :\ \mathbb{R} \to \mathbb{R}$?

b. Considérense las ecuaciones

$$\begin{cases} x^3 + y^{40} & = 2 \\ xz + y^2 + y & = 3. \end{cases}$$

Muéstrese que se pueden despejar $y(x)$, $z(x)$ cerca de $(x, y, z) = (1, 1, 1)$. Calcúlese dy/dx en $x = 1$.

c. Sea $\varphi : [0, 1] \to [0, 1]$ continua. Demuéstrese que φ tiene un punto fijo.

d. Sea $F :\ \mathbb{R}^n \to \mathbb{R}^n$ de clase C^1 y con jacobiano distinto de cero en todo punto. Demuéstrese que $F(\mathbb{R}^n)$ es abierto.

e. Sea $f :\ \mathbb{R}^2 \to \mathbb{R}$ continua. Muéstrese que f no es inyectiva. [Sugerencia: si f fuera inyectiva, entonces las imágenes de los ejes X e Y serían intervalos en $\mathbb{R}$.]

61. **a.** Evalúese $\int_A e^{-x^2-y^2}\, dxdy$, donde $A = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$.

b. Evalúese $\int_B x\,dxdy$, donde B es la región del plano acotada por $x = 0$, $y = 0$ y $x + y = 1$.

c. Enúnciese una versión del teorema de Fubini.

d. Úsese el apartado **c** para escribir una fórmula para $\int_A f(x, y)\,dxdy$, donde $A = \{(x, y) \in \mathbb{R}^2 \mid c \leq y \leq d \text{ y } \psi(y) \leq x \leq \varphi(y)\}$. En este caso, ψ y φ son funciones suaves en el intervalo $[c, d]$ tales que $\psi(y) \leq \varphi(y)$ para todo $y \in [c, d]$. Dibújese la región.

e. Sea $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ de clase C^1 y biyectiva, con $J\varphi \neq 0$. Supóngase que $\int_A dxdy = \int_{\varphi(A)} dxdy$ para todo disco abierto A. Demuéstrese que $J\varphi = 1$.

62. a. Sea V un espacio con producto interno y $\varphi_0, \varphi_1, \varphi_2, \ldots$ una sucesión ortonormal en V. Escríbase la serie de Fourier de $f \in V$ con respecto a $\{\varphi_i\}$.

b. Explíquese la forma en que el apartado **a** se relaciona con la fórmula

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + \sum_{n=1}^{\infty} b_n \operatorname{sen} nx$$

donde

$$\begin{aligned} a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx\,dx \quad k = 0, 1, 2, \ldots \\ b_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\operatorname{sen} kx\,dx \quad k = 1, 2, \ldots. \end{aligned}$$

c. Calcúlese la serie de Fourier de $f(x) = x$, $-\pi < x < \pi$.

d. ¿Cuál es el límite puntual de la serie del apartado **c**? ¿Converge dicha serie en media? Razónese la respuesta.

e. Suponiendo los teoremas de completitud del texto, demuéstrese que $\{\operatorname{sen} nx \mid n = 1, 2, 3, \ldots\}$ es completo en el espacio de las funciones de cuadrado integrable en $[0, \pi]$.

Los ejercicios 63–66 se basan en los capítulos 1–7.

63. Sea $B = D(0, 1)$ la bola unidad abierta en $\mathbb{R}^n$ con centro en 0 y sea $f : B \to \mathbb{R}$ continua. Supóngase que existe una transformación continua $g : B \to L(\mathbb{R}^n, \mathbb{R})$ (las transformaciones lineales de $\mathbb{R}^n$ a $\mathbb{R}$), tal que para cada par de puntos x, y en B, $f(y) - f(x) = \int_0^1 [g(ty + (1-t)x))(y - x)]\,dt$. Muéstrese que f es de clase C^1 y que $Df = g$.

64. Sea $D = \{x \in \mathbb{R}^2 \mid \|x\| \leq 1\}$ el disco unidad cerrado con centro en el origen de $\mathbb{R}^2$. Para cada entero $n > 0$, sea $f_n : D \to \mathbb{R}$ continua y supóngase que $f : D \to \mathbb{R}$ es una función continua tal que la sucesión $\langle f_n\rangle_1^\infty$ converge uniformemente a f en D. ¿Es equicontinuo o acotado (o ambas cosas) el conjunto de funciones $\{f_n \mid n = 1, 2, 3, \ldots\}$? Justifíquese la respuesta.

65. Para cada par de funciones $f, g : [0, 1] \to \mathbb{R}$, defínase $f \vee g : [0, 1] \to \mathbb{R}$ como $(f \vee g)(x) = \text{máx}\{f(x), g(x)\}$.

a. Si f y g son continuas, muéstrese que $f \vee g$ es continua.

b. Defínase una transformación $\psi : \mathcal{C}([0, 1], \mathbb{R}) \times \mathcal{C}([0, 1], \mathbb{R}) \to \mathcal{C}([0, 1], \mathbb{R})$ como $\psi(f, g) = f \vee g$. Muéstrese que ψ es continua (con respecto a la norma del supremo en el espacio $\mathcal{C}([0, 1], \mathbb{R})$ de todas las funciones continuas de valores reales en $[0, 1]$).

66. Defínase una transformación $\alpha : \mathcal{C}([0, 1], \mathbb{R}) \to \mathcal{C}([0, 1], \mathbb{R})$ como sigue: para cada $f \in \mathcal{C}([0, 1], \mathbb{R})$, defínase $\alpha(f) \in \mathcal{C}([0, 1], \mathbb{R})$ como $(\alpha(f))(x) = \int_0^x f(t)\, dt$.

a. Muéstrese que α es una transformación continua de $\mathcal{C}([0, 1], \mathbb{R})$ en $\mathcal{C}([0, 1], \mathbb{R})$ con respecto a la norma del supremo en $\mathcal{C}([0, 1], \mathbb{R})$.

b. ¿Es α una transformación lineal compacta? Justifíquese la respuesta (una transformación lineal es compacta si la clausura de la imagen de la bola unidad es compacta).

Apéndice B

Referencias y sugerencias para estudios posteriores

Este apéndice contiene referencias del texto y varias sugerencias para un estudio más a fondo de la materia. Se agrupan por temas, y hemos dado unas cuantas observaciones acerca del nivel y material de las selecciones.

El número de libros de cálculo avanzado y de introducción al análisis es extraordinario. A pesar de los muchísimos textos recientes, algunos de los libros más antiguos siguen siendo los mejores. Algunos de nuestra preferencia son:

[1] Carslaw, H. S., *Theory of Fourier's Series and Integrals*, 3ª ed., Nueva York, Dover, 1930.

[2] Hardy, G. H., *Pure Mathematics*, 9ª ed., Nueva York, Cambridge Univ. Press, 1947.

[3] Hobson, E. W., *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, Cambridge, Ingl., Cambridge Univ. Press, 1921.

[4] Titchmarsh, E. C., *Theory of Fourier Integrals*, Londres, Oxford Univ. Press, 1937.

[5] Whittaker, E. T., y Watson, G. N., *A Course of Modern Analysis*, Cambridge, Ingl., Cambridge Univ. Press, 1926.

De los textos más recientes que están aproximadamente al mismo nivel que los mencionados, los siguientes han tenido gran aceptación. De estos, [6, 7, 8, 9, 12, 13, 15, 16] son bastante clásicos, mientras que [10, 11, 14] tienden a ser un poco más abstractos.

[6] Apostol, T. M., *Mathematical Analysis*, 2ª ed., Reading, Mass., Addison-Wesley, 1974.

[7] Bartle, R. G., *The Elements of Real Analysis*, 2ª ed., Nueva York, Wiley, 1976.

[8] Buck, R. C., *Advanced Calculus*, 2ª ed., Nueva York, McGraw-Hill, 1965.

[9] Graves, L. M., *Theory of Functions of Real Variables*, 2ª ed., Nueva York, McGraw-Hill, 1956.

[10] Lang, S., *Analysis I*, Reading, Mass., Addison-Wesley, 1968.

[11] Loomis, L. H. y Sternberg, S., *Advanced Calculus*, Reading, Mass., Addison-Wesley, 1968.

[12] Olmsted, J. M. A., *Advanced Calculus*, Nueva York, Appleton-Century-Crofts, 1961.

[13] _____, *Real Variables*, Nueva York, Appleton-Century-Crofts, 1956.

[14] Rosenlicht, M., *Introduction to Analysis*, Glenview, Ill., Scott, Foresman and Co., 1968.

[15] Rudin, W., *Principles of Mathematical Analysis*, 3ª ed., Nueva York, McGraw-Hill, 1976.

[16] Widder, D. V., *Advanced Calculus*, 2ª ed., Englewood Cliffs, N.J., Prentice-Hall, 1965.

Para mayor información acerca de los fundamentos de la teoría de conjuntos, consúltese [17, 18, 19]. En [17] no hay mucho material sobre lógica y los axiomas de teoría de conjuntos, pero todos los resultados de la teoría de conjuntos que son de importancia práctica en el curso se presentan de forma concisa. Además, [17] desarrolla ampliamente el cálculo diferencial abstracto (véanse nuestros capítulos 6 y 7) en el contexto de los espacios de Banach. Los axiomas de conjuntos de la introducción se tomaron textualmente de [18]. [19] es un texto más moderno diseñado para presentar al estudiante la teoría básica de conjuntos e iniciar el estudio de la materia en sí.

[17] Dieudonné, Jean, *Foundations of Modern Analysis*, Nueva Jersey, Prentice-Hall, 1966.

[18] Halmos, Paul R., *Naive Set Theory*, Nueva York, Springer-Verlag, 1960.

[19] Roitman, J., *Introduction to Modern Set Theory*, Nueva York, Wiley, 1990.

Las siguientes son algunas referencias generales para un trabajo más avanzado en análisis real, incluyendo la integración de Lebesgue y el análisis abstracto en espacios generales de Banach y de Hilbert.

[20] Burkhill, J. C., *The Lebesgue Integral*, Cambridge, Ingl., Cambridge Univ. Press, 1951.

[21] Halmos, P. R., *Measure Theory*, Nueva York, Springer-Verlag, 1950.

[22] Hewitt, E. y Stromberg, K., *Real and Abstract Analysis*, Nueva York, Springer-Verlag, 1969.

[23] Gleason, A. M., *Fundamental of Abstract Analysis*, Reading, Mass., Addison-Wesley, 1966.

[24] Lang, S., *Analysis II*, Reading, Mass., Addison-Wesley, 1969.

[25] Royden, H. L., *Real Analysis*, 3ª ed., Nueva York, Macmillan. 1989.

[26] Rudin, W., *Real and Complex Analysis*, 3ª ed., Nueva York, McGraw-Hill, 1987.

[27] _____, *Functional Analysis*, 2ª ed., Nueva York, McGraw-Hill, 1991.

[28] Simmons, G., *Introduction to Topology and Modern Analysis*, Nueva York, McGraw-Hill, 1963.

Un libro útil para encontrar contraejemplos a teoremas con hipótesis de menos es [29].

[29] Gelbaum, B. R. y Olmsted, J.M.H., *Counterexamples in Analysis*, San Francisco, Holden Day, 1964.

Nuestro texto estudia un poco de series. Referencias clásicas sobre el tema son [30, 31].

[30] Hardy, G. H., *Divergent Series*, Londres, Oxford Univ. Press, 1949.

[31] Knopp, K., *Theory and Application of Infinite Series*, Nueva York, Dover, 1951.

Para quien desee estudiar con más detalle la teoría de distribuciones, puede consultar los siguientes títulos, además de [27].

[32] Gelfand, I.M. y Shilov, G.E., *Generalized Functions*, Nueva York, Academic Press, 1964.

[33] Schwartz, L., *Théorie des distributions*, París, Hermann. 1966.

[34] Zemanian, A., *Distribution Theory and Transform Analysis*, Nueva York, McGraw-Hill, 1965.

Los siguientes títulos desarrollan la teoría de ecuaciones diferenciales y ecuaciones integrales. De éstos, [35] y [36] son tratados amplios.

[35] Coddington, E. A. y Levinson, N., *Theory of Ordinary Differential Equations*, Nueva York, McGraw-Hill, 1955.

[36] Hartman, P., *Ordinary Differential Equations*, Nueva York, Wiley, 1964.

[37] Hurewicz, W., *Lectures on Ordinary Differential Equations*, Nueva York, Dover, 1964.

[38] Roxin, E. O., *Ordinary Differential Equations*. Belmont, Calif., Wadsworth, 1972.

[39] Widom, H., *Lectures on Integral Equations*, Nueva York, Van Nostrand Mathematical Studies No. 17, 1969.

El cálculo avanzado se puede aplicar con elegancia para estudiar problemas en geometría y análisis vectorial, incluyendo la geometría fractal. Además de [10, 11, 24], consúltese:

[40] Barnsley, M., *Fractals Everywhere*, San Diego, Academic Press, 1988.

[41] Devaney, R. L., *An Introduction to Chaotic Dynamical Systems*, 2ª ed., Redwood City, Calif., Addison-Wesley, 1989.

[42] Fleming, W., *Functions of Several Variables*, Reading, Mass., Addison-Wesley, 1965.

[43] Munkres, James R., *Analysis on Manifolds*, Redwood City, Calif., Addison-Wesley, 1991.

[44] Peitgen, H.-O., Jürgens, H. y Saupe, D., *Chaos and Fractals*; *New Frontiers of Science*, Nueva York, Springer-Verlag, 1992.

[45] Spivak, M., *Calculus on Manifolds*, Nueva York, Benjamin, 1965.

Hemos citado varios textos que versan sobre las series de Fourier y el análisis: [1, 3, 4, 22, 24, 26, 33, 34]. Otros textos, un poco más avanzados, son:

[46] Körner, T.W., *Fourier Analysis*, Cambridge, Ingl., Cambridge Univ. Press, 1988.

[47] Stein, M. y Weiss, G., *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton, N. J., Princeton Univ. Press, 1971.

[48] Widom, H., *Lectures on Measures and Integration*, Nueva York, Van Nostrand Mathematical Studies No. 20, 1969.

[49] Zygmund, A., *Trigonometric Series*, 2ª ed., Cambridge, Ingl., Cambridge Univ. Press, 1959.

Nuestro capítulo acerca del análisis de Fourier dio una introducción a las ecuaciones diferenciales en derivadas parciales. Se puede encontrar más información en los siguientes textos, de los cuales, los dos últimos usan la teoría de distribuciones, y [53] es más avanzado.

[50] Churchill, R. V. y Brown, J. W., *Fourier Series and Boundary Value Problems*, 4ª ed., Nueva York, McGraw-Hill, 1987.

[51] Courant, R., y Hilbert, D., *Methods of Mathematical Physics* (2 volúmenes), Nueva York, Wiley-Interscience, 1962.

[52] Duff, G. F. D. y Naylor, D., *Differential Equations of Applied Mathematics*, Nueva York, Wiley, 1966.

[53] Sobolev, S. L., *Applications of Functional Analysis in Mathematical Physics*, Providence, R.I., American Mathematical Society Translations, vol. 7, 1963.

Las referencias que siguen a continuación tratan acerca de la mecánica cuántica. [56] es un texto clásico elemental, mientras que [54, 55] son más avanzados y orientados a los matemáticos.

[54] Jauch, J. M., *Foundations of Quantum Mechanics*, Reading, Mass., Addison-Wesley, 1968.

[55] Mackey, G. W., *The Mathematical Foundations of Quantum Mechanics*, Nueva York, Benjamin, 1963.

[56] Merzbacher, E., *Quantum Mechanics*, 2ª ed., Nueva York, Wiley, 1970.

Existen varios temas importantes en el análisis clásico que no hemos estudiado. Por ejemplo, podríamos haber estudiado la función gamma, según [16] o [57]. Este tema se estudia también con frecuencia en cursos de variable compleja.

[57] Artin, E., *The Gamma Function*, Nueva York, Holt, Rinehart and Winston, 1964.

Existen numerosos textos excelentes de autores que no son de habla inglesa; por ejemplo:

[58] Bourbaki, N., *Éléments de Mathématique; Fonctions d'une variable réele*, París, Hermann, 1961.

[59] Dieudonné, J., *Calcul Infinitésimal*, París, Hermann, 1971.

El tratamiento riguroso del análisis elemental no evolucionó de forma rápida ni fluida. Los creadores de esta área de las matemáticas encontraron muchos obstáculos y callejones sin salida en el camino antes de desembocar en sus brillantes ideas. La apreciación de su historia es importante para la educación del estudiante en matemáticas. Los textos recomendables son:

[60] Grabiner, J. V., *The Origins of Cauchy's Rigorous Calculus*, Cambridge, Mass., The MIT Press, 1981.

[61] Kline, M., *Mathematical Thought from Ancient to Modern Times*, Nueva York, Oxford Univ. Press, 1972.

Apéndice C

Respuestas y sugerencias a los ejercicios impares

En este apéndice damos las respuestas o sugerencias de muchos de los ejercicios impares. Pocas veces se tratara de las soluciones completas. Hay que realizar los cálculos y dar ciertos argumentos para apoyar las respuestas o rellenar los huecos y seguir las pistas y sugerencias.

Introducción: Conjuntos y funciones

1. **a.** $f(A) = \{1\}, f^{-1}(B) = S$.

b. $f(A) = A, f^{-1}(B) = B$.

c. $f(A) = \{1, 0, -1\}, f^{-1}(B) = \{x \mid x \leq 0\}$.

3. En cada caso, úsense las definiciones. Por ejemplo, para **a**, $x \in f^{-1}(C_1 \cup C_2)$ sii $f(x) \in C_1 \cup C_2$ sii $(f(x) \in C_1$ o $f(x) \in C_2)$ sii $x \in f^{-1}(C_1) \cup f^{-1}(C_2)$. No necesariamente ocurre la igualdad en **d**.

5. Por ejemplo, supóngase que $f(D_1 \cap D_2) = f(D_1) \cap f(D_2)$. Si $x_1 \neq x_2$ y $f(x_1) = f(x_2) = y$, sean $D_1 = \{x_1\}$ y $D_2 = \{x_2\}$ para deducir una contradicción. Así, f es inyectiva. Algunos posibles equivalentes a la "suprayectiva" son: para todo $y \in T, f^{-1}(\{y\}) \neq \varnothing$ o, para todo $R \subset T, R = f(f^{-1}(R))$. Son posibles otras respuestas.

7. Un subconjunto B de A se determina viendo si para cada $x \in A$, $x \in B$ o $x \notin B$. Aplíquese un argumento estándar para contar una sucesión de elecciones o bien úsese la inducción.

9. Si $x \in \bigcup \mathcal{A}$, entonces $x \in A$ para algún $A \in \mathcal{A}$. Como $\mathcal{A} \subset \mathcal{B}$, $A \in \mathcal{B}$ y entonces $x \in \bigcup \mathcal{B}$. Así, $\bigcup \mathcal{A} \subset \bigcup \mathcal{B}$. La segunda parte es análoga.

11. $x \in (g \circ f)^{-1}(C) \iff (g \circ f)(x) \in C \iff g(f(x)) \in C \iff f(x) \in g^{-1}(C) \iff x \in f^{-1}(g^{-1}(C))$.

13. $A, B \neq \varnothing \iff (\exists a \in A \text{ y } \exists b \in B) \iff (a, b) \in A \times B$, y entonces $A \times B \neq \varnothing$.

15. **a.** Si f es inyectiva, sea $g(y) = x$ si $f(x) = y$ y sea $g(y)$ cualquier cosa si no existe tal x.

b. Si f es suprayectiva e $y \in T$, elíjase algún x tal que $f(x) = y$ y sea $h(y) = x$ (obsérvese el uso del axioma de elección).

17. Sea $T: \mathbb{R}^n \to \mathbb{R}^m$ la transformación lineal asociada a la matriz A. Por álgebra lineal sabemos que T es suprayectiva; es decir, $Tx = y$ se puede resolver para $x \in \mathbb{R}^n$ para cada $y \in \mathbb{R}^m$ si y sólo si T tiene rango m. Si B existe, T es suprayectiva por el ejercicio **15b.** Recíprocamente, si T es suprayectiva, elíjase un subespacio $W \subset \mathbb{R}^n$ de dimensión m complementario a $\ker(T)$. Entonces $T \mid W : W \to \mathbb{R}^m$ es un isomorfismo. Sea U su inversa. Entonces $T \circ U =$ identidad. Si B es la matriz de U, obtenemos $AB =$ identidad.

Capítulo 1: La recta real y el espacio euclídeo

1.1 Cuerpos ordenados y los sistemas numéricos

1. $(a+b)^2 = (a+b)(a+b) = (a+b)a + (a+b)b = a^2 + ba + ab + b^2$ (por la propiedad distributiva, dos veces). Úsese ahora la conmutatividad.

3. Desarróllese $(a-b)^3 > 0$ y reordénese.

5. Sea $\mathbb{F} = \{0, 1, 2\}$ con aritmética módulo 3. Por ejemplo, $2 \cdot 2 = 1$ y $1 + 2 = 0$. Para mostrar que no puede ser ordenado, obténgase una contradicción de (por ejemplo) $1 > 0$, de modo que $1 + 1 = 2 > 0$ por lo que $1 + 2 = 0 > 0$.

1.2 La completitud y el sistema de los números reales

1. **a.** Tómense límites cuando $n \to \infty$ en $x_n^2 = 2 + x_{n-1}$.

b. 2.

3. 0.

5. De una sucesión monótona (no trivial) $\langle x_n \rangle_1^\infty$, extráigase una subsucesión que sea estrictamente monótona.

1.3 Supremos

1. $\sup(S) = 1$; S no está acotado inferiormente.

3. $\sup(Q)$ es una cota superior de P, de modo que $\sup(Q) \geq \sup(P)$.

5. $\sup(S) = 1$.

1.4 Sucesiones de Cauchy

1. $|x_n - x_{n+k}| \leq \sum_{i=n}^{n+k-1} |x_i - x_{i+1}| \leq \sum_{i=n}^{\infty} (1/3^i) \leq 1/3^{n-1}$. Así, para todo $\varepsilon > 0$, podemos elegir N tal que $1/3^{n-1} < \varepsilon$. Obtenemos una sucesión de Cauchy, por lo que $\langle x_n \rangle_1^\infty$ converge.

3. Una posibilidad es $1, 0, -1, 1, 0, -1, 1, 0, -1, \ldots$.

5. Falso.

1.5 Puntos límite; lím inf y lím sup

1. $\text{lím inf}(x_n) = 2$; $\text{lím sup}(x_n) = 4$.

3. Úsese **1.5.5** para mostrar que existen puntos $x_{N(n)}$ a una distancia menor que $1/n$ de A. Asegúrese el lector de obtener una subsucesión.

5. Falso.

1.6 El espacio euclídeo

1. Si la igualdad es válida en la desigualdad de Cauchy-Schwarz o la desigualdad triangular, entonces x e y son paralelos. Desarróllese $\|x + y\|^2 = (\|x\| + \|y\|)^2$ para obtener $\|x\| \cdot \|y\| = \langle x, y \rangle = \|x\| \cdot \|y\| \cdot \cos(\theta)$. Úsese esto directamente o tómense $u = y/\|y\|$ y $z = x - \langle x, u \rangle u$. Verifíquese que $\langle z, u \rangle = 0$ y que $\|x\|^2 = \|z\|^2 + |\langle x, u \rangle|^2 = \|x\|^2$. Conclúyase que $z = 0$ y $x = (\langle x, y \rangle / \|y\|^2) y$.

3. $\{\lambda \cdot (-2, 0, 3) \in \mathbb{R}^3 \mid \lambda \in \mathbb{R}\}$.

5. $x = (y + 1)/2 = (z + 2)/3$ o $P(t) = (1 + t, 1 + 2t, 1 + 3t)$. Esta recta no es un subespacio lineal, pues (0, 0, 0) no está en ella.

1.7 Normas, productos internos y métricas

1. La norma del supremo es 1. La de **1.7.7** es $1/\sqrt{3}$.

3. $\sqrt{\langle f,f \rangle} = 1$ $\sqrt{\langle g,g \rangle} = 1/\sqrt{3}$, y $\langle f,g \rangle = 1/2$. Así, $|\langle f,g \rangle| \leq \sqrt{\langle f,f \rangle} \cdot \sqrt{\langle g,g \rangle}$ es verdadero.

5. Para $f(x) = x$, $\|f\|_\infty = 1$, pero $\langle f, f \rangle^{1/2} = 1/\sqrt{3}$. Así, estas normas son diferentes.

1.8 Los números complejos

1. **a.** $6 + 4i$.

b. $(11/5) - 2i$.

c. $(3/2) - (5/2)i$.

3. No, no siempre.

5. No; cierto sii z es real.

7. **a.** Úsense las identidades trigonométricas y las propiedades de los exponentes reales.

b. $|e^z| = e^{Re(z)} \neq 0$; así, $e^z \neq 0$.

c. $\cos^2 \theta + \operatorname{sen}^2 \theta = 1$

d. Úsese la inducción.

9. $|z| < 1$; entonces $z_n \to 0$ cuando $n \to \infty$.

Ejercicios del capítulo 1 (al final del capítulo)

1. **a.** $\sup(S) = \sqrt{5}$; $\inf(S) = -\sqrt{5}$.

b. Ni $\sup(S)$ ni $\inf(S)$ existen (excepto como $\pm\infty$).

c. $\sup(S) = 1$; $\inf(S) = 0$.

d. $\sup(S) = 0$; $\inf(S) = -1$.

e. $\sup(S) = 1/3$; $\inf(S) = 0.3$.

3. **a.** Supóngase que $x > 0$ y considérese $\varepsilon = x/2$.

b. Sea $x = \text{mín}\{\varepsilon/2, 1/2\}$.

5. Si $\varepsilon > 0$, úsese **1.3.2** para obtener N tal que $\sup(S) - \varepsilon < x_N \leq \sup(S)$. Si $n \geq N$, entonces $\sup(S) - \varepsilon < x_N \leq x_n \leq \sup(S)$, de modo que $|x_n - \sup(S)| < \varepsilon$. Si $\langle y_n \rangle_1^\infty$ es monótona decreciente y $R = \{y_1, y_2, y_3, \ldots\}$, entonces $y_n \to \inf(R)$.

7. Como $\sup(A) + \sup(B)$ es una cota superior de $A + B$ (¿por qué?), $\sup(A + B) \leq \sup(A) + \sup(B)$. A continuación, muéstrese que los elementos de $A + B$ están arbitrariamente cerca de $\sup(A) + \sup(B)$.

9. Úsese $y_n \leq |x_n|$. La igualdad no necesariamente es cierta. Por ejemplo, sea $x_n = (-1)^{n+1}$. Entonces $\text{lím sup } y_n = -1$ y $\text{lím sup } |x_n| = 1$; $\text{lím inf } y_n \geq \text{lím inf}(-|x_n|)$.

11. **i** $\Rightarrow$ completitud: sea $\langle x_n \rangle_1^\infty$ creciente y acotada superiormente. Por **i**, $\langle x_n \rangle_1^\infty$ tiene un *supremo*, digamos S. Verifíquese que $x_n \to S$. Véase el ejercicio 5. La demostración de que **ii** $\Rightarrow$ completitud se hace mostrando que $x_n \to -\inf\{-x_n\}$.

13. El espacio generado por $\{(2, 1, -2, 0), (2, 1, 0, -2)\}$.

15. Muéstrese primero que $d(x_n, x_{n+1}) \leq d(x_1, x_2)/2^{n-1}$ y luego que $d(x_n, x_{n+k}) \leq d(x_1, x_2)/2^{n-2}$.

17. Si L es el conjunto de cotas inferiores de S e $y \in L$, entonces $y \leq \inf(S)$ (¿por qué?). Así, $\sup(L) \leq \inf(S)$. Pero $\inf(S) \in L$, por lo que $\inf(S) \leq \sup(L)$.

19. $y_1 = y_2 = y_3 = 1/2$.

21. Si $x \in S$, sea $x_k = x$. En caso contrario, $x_k < x$ para todo k. Sea $y_1 = x_1$, y_2 el primero de los valores $x_2, x_3, \ldots$ tal que $y_1 \leq y_2$, etcétera.

23. Como 0 es una cota inferior de P, $0 \leq \inf(P)$. Úsese la condición dada para descartar $0 < \inf(P)$.

25. **a.** Si $x \in P$, entonces existe $y \in Q$ tal que $x \leq y \leq \sup(Q)$. Así, $\sup(Q)$ es una cota superior de P.

b. No: considérese $P = [0, 1]$ y $Q = [-1, 1]$. Entonces $P \leq Q$ pero $\inf(P) > \inf(Q)$.

c. No: considérese P = irracionales y Q = racionales. He aquí otro ejemplo: $P = \{1\}$ y $Q = \{0, 1\}$. Entonces $P \leq Q$, pero $\inf(P) > \inf(Q)$. Además, $Q \leq P$, pero $P \neq Q$.

27. $\inf(B) = \sqrt{2}$.

29. Sea $y = \sup\{u \in \mathbb{R} \mid u^2 \leq x\}$. Obténganse contradicciones de $y^2 < x$ y de $y^2 > x$.

31. Cierto.

33. **a.** Úsese la forma ∞/∞ de la regla de l'Hôpital.

b. Tómense logaritmos y úsese **a.**

35. $(3 - 8i)^4/(1 - i)^{10}$.

37. $z = \pm(2 - i)$.

39. **a.** $\text{Re}\left(1/z^2\right) = (x^2 - y^2)/(x^2 + y^2)$; $\text{Im}\left(1/z^2\right) = -2xy/(x^2 + y^2)$.

b. $\text{Re}\left(1/(3z + 2)\right) = (3x + 2)/((3x + 2)^2 + 9y^2)$; $\text{Im}\left(1/(3z + 2)\right) = -3y/((3x + 2)^2 + 9y^2)$

41. Si $z = x + iy$, entonces $\text{Re}(iz) = \text{Re}(ix - y) = -y = -\text{Im}(z)$, e $\text{Im}(z) = \text{Im}(ix - y) = x = \text{Re}(z)$.

43. Úsese $|z|^2 = z \cdot \bar{z}$ y $\overline{z + w} = \bar{z} + \bar{w}$ para desarrollar $|a - b|^2 + |a + b|^2$.

45. $|z|^n/n > R \Leftrightarrow \log |z| > (\log n + \log R)/n$. El miembro izquierdo es positivo y el miembro derecho tiende a 0 por la regla de l'Hôpital.

Capítulo 2: La topología del espacio euclídeo

2.1 Conjuntos abiertos

1. Si $v = (a, b) \neq (0, 0)$, entonces $v \in U = D(v, \|v\|) \subset \mathbb{R}^2 \setminus \{(0, 0)\}$ (proporciónense los detalles).

3. Si $v_0 = (x_0, y_0) \in B$, existe $r > 0$ tal que $|x - x_0| < r \Rightarrow x \in A$. (¿Por qué?). Muéstrese que $\|v - v_0\| < r \Rightarrow v \in B$.

5. No en general; sí si B es abierto.

2.2 Interior de un conjunto

1. $\text{int}(S) = \{(x, y) \in \mathbb{R}^2 \mid xy > 1\}$.

3. Sí.

5. Más generalmente, muéstrese que si U es abierto y $U \subset A$, entonces $U \subset \text{int}(A)$ (véase el ejercicio 3 al final del capítulo).

2.3 Conjuntos cerrados

1. Sí.

3. Si $S = \{x_1, \ldots, x_k\}$ y $v_0 \in \mathbb{R}^n \setminus S$, sea $r = (1/2)\ \text{mín}\{d(v_0, x_1), \ldots, d(v_0, x_k)\}$. Muéstrese que $\|v - v_0\| < r \Rightarrow v \in \mathbb{R}^n \setminus S$.

5. No.

2.4 Puntos de acumulación

1. $\{(x, y) \in \mathbb{R}^2 \mid y = 0 \text{ y } 0 \leq x \leq 1\}$.

3. **a.** $\varnothing$.

b. $\mathbb{R}^2$.

c. Eje X.

d. $\{(1/n, 0) \in \mathbb{R}^2 \mid n \in \mathbb{Z}, n \neq 0\} \cup \{0\}$.

5. A es la parábola $y = 4 - (x + 1)^2$. Si (a, b) es un punto de acumulación, entonces existen (x_k, y_k) en ella tales que $x_k \to a$ e $y_k \to b$. Por continuidad, $b = 4 - (1 + a)^2$. Así, $(a, b) \in A$.

2.5 Clausura de un conjunto

1. $\text{cl}(S) = \{(x, y) \in \mathbb{R}^2 \mid x \geq y^2\}$.

3. $\mathbb{R}^2$.

5. El conjunto $B(x, r) = M \setminus \{y \mid D(y, x) > r\}$ es cerrado y $A \subset B(x, r)$. Así, $\text{cl}(A) \subset b(x, r)$.

2.6 Frontera de un conjunto

1. $\partial A = \{0\} \cup A$.

3. $\partial A = \{(x, y) \in \mathbb{R}^2 \mid x = y\}$.

5. Sí.

2.7 Sucesiones

1. (0, 0).
3. Úsese el teorema **2.7.6 ii.**
5. $\mathrm{cl}(S) = \{x \in \mathbb{R} \mid x^2 \leq 2\} = [-\sqrt{2}, \sqrt{2}]$.

2.8 Completitud

1. Si $x_n \in N$ y $x_n \to x$, entonces $\langle x_n \rangle_1^\infty$ es una sucesión de Cauchy. Como N es completo, esta sucesión tiene un límite en $N \subset M$. Como los límites en M son únicos, este límite debe ser x. Esto muestra que N es cerrado (¿por qué?).
3. Sí, toda sucesión de Cauchy es tarde o temprano constante.
5. Si $\langle x_n \rangle_1^\infty$ es una sucesión de Cauchy, ésta está acotada. Tiene un punto de acumulación y por lo tanto una subsucesión convergente. Así, la sucesión completa converge (¿por qué?).

2.9 Series de números reales y de vectores

1. Converge.
3. Sea $\varepsilon > 0$. Por **2.9.2**, existe N tal que $\|x_k\| < \varepsilon$ para todo $k \geq N$. Así, $x_k \to 0$.
5. No converge.

Ejercicios del capítulo 2 (al final del capítulo)

1. **a.** Abierto.
   b. Cerrado.
   c. $[-1, 0]$ es cerrado.
   d. Abierto y cerrado.
   e. Cerrado.
   f. Ni abierto ni cerrado.
   g. Ni abierto ni cerrado.
   h. Cerrado.

3. Cada x en U satisface la definición de un punto interior de A. Para conjuntos cerrados: si B es cerrado y $B \subset A$, entonces $B \subset \text{cl}(A)$. O bien: si $A \subset C$ y C es cerrado, entonces $\text{cl}(A) \subset C$.

5. Existe un conjunto abierto U tal que $x \in U \subset A$. Úsese ahora el hecho de que U es abierto.

7. No es cierto para todo subconjunto de M. Inténtese con $U = [0, 1] \subset \mathbb{R}$.

9. Úsese **2.6.2** para obtener **a** y luego úsese **a** para obtener **b.**

11. Considérese $d(x_m, x) \leq \varepsilon/2 < \varepsilon$.

13. Compárese con el ejercicio 9.

15. **a.** Úsese $M \setminus (M \setminus A) = A$ y la definición de frontera.

b. Úsese el hecho de que ∂A es cerrado (¿por qué?).

17. La serie $\sum_m \|(\text{sen}\, m)x_m\|$ converge, por comparación con $\sum_m \|x_m\|$ (¿por qué?). Así, $\sum_m x_m$ converge, por **2.9.3.**

19. **a.** Un punto límite o bien pertenece a A o bien es un punto de acumulación de A. Podría ser un punto aislado de A; un punto de acumulación no puede serlo.

21. Cada vecindad abierta U de 0 contiene un disco $D(0, \varepsilon)$.

23. Si V es la unión de todos los subconjuntos abiertos de A, entonces V es abierto y está contenido en A. Así, $V \subset \text{int}(A)$ (véase el ejercicio 3). Si $x \in \text{int}(A)$, existe un abierto U tal que $x \in U \subset A$, de modo que $x \in V$.

25. Como los discos de radio ε son abiertos, también lo es cualquier unión de ellos. Recíprocamente, si A es abierto y $x \in A$, existe $\varepsilon_x > 0$ tal que $x \in D(x, \varepsilon_x) \subset A$. Así, $A = \bigcup_{x \in A} D(x, \varepsilon_x)$.

27. Elíjase b_n tal que $b_n = |b_n| < \varepsilon/2^n$.

29. Sí.

31. Lo que debe hacerse es mostrar que un punto de acumulación de A' debe ser un punto de acumulación de A. $(A')'$ no tiene por qué ser igual a A'. Considérese $A = \{1/2, 1/3, \ldots\}$.

33. Muéstrese que $\alpha_n = s_{n+1} - s_n$ es creciente y que está acotada superiormente. Úsese la serie telescópica $s_n = s_0 + \sum_{k=1}^{n-1} \alpha_k$ para mostrar que el límite debe ser 0.

35. Sugerencia: cada uno de los conjuntos debe contener un número racional.

37. $[A \cap \text{cl}\,(M \setminus A)] \cup [\text{cl}\,(A) \setminus A] = [A \cup (\text{cl}\,(A) \setminus A)] \cap [\text{cl}\,(M \setminus A) \cup (\text{cl}\,(A) \setminus A)]$ $= \text{cl}\,(A) \cap [\text{cl}\,(M \setminus A) \cup (\text{cl}(A) \setminus A)]$. ¿Por qué es esto ∂A?

39. Muéstrese primero que $\sup(S) - \inf(S)$ es una cota superior para $\{x - y \mid x \in S$ e $y \in S\}$. Después, considérese x e y en S muy cercanos a $\sup(S)$ y a $\inf(S)$.

41. Sea U una vecindad de x. Existe n tal que $x \in \text{cl}\,(A_n) \setminus A_n$ (¿por qué?). Así, $U \cap (A_n \setminus \{x\}) \neq \varnothing$.

43. $(1 + \sqrt{13})/2$.

45. Necesitamos que $x^{\log x}/e^x \to 0$. Háganse algunos cálculos y úsese la regla de l'Hôpital.

47. Compárese la gráfica de $1/x$ con los rectángulos inscritos y circunscritos para mostrar que la sucesión $\gamma_n = 1 + (1/2) + \cdots + (1/n) - \log n$ es decreciente y que está acotada inferiormente por 0.

49. Muéstrese que $a_n = O(n^{-A})$ analizando $P_n = \Pi_{k=1}^{n} (1 - A/k)$; úsese el ejercicio 47 para establecer que $P_n = -A \log n + O(1)$.

51. **d.** $1/e$.

53. 14.

Capítulo 3: Conjuntos compactos y conexos

3.1 Compacidad

1. Sea $\langle x_n \rangle_1^\infty$ una sucesión en A. Si $\{x_n\}$ es un conjunto finito, alguna entrada debe repetirse una infinidad de veces. Úsese ésta para obtener una subsucesión convergente. Si es un conjunto infinito, entonces un punto de acumulación debe ser el límite de una subsucesión (es preciso tener cuidado, pues esto es más complejo de lo que aparenta: necesitamos el límite de una subsucesión, no de cualquier sucesión).

3. $\text{cl}\,(A)$ es cerrado. Para mostrar que es totalmente acotado, sea $\varepsilon > 0$ y recubramos A con una cantidad finita de bolas de radio $\varepsilon/2$. Muéstrese que $\text{cl}\,(A)$ queda recubierta por las bolas correspondientes de radio ε.

5. Una sucesión converge sii tarde o temprano se vuelve constante. Esto no contradice el ejercicio 4, pues las entradas de tal sucesión convergente forman un conjunto finito.

3.2 Teorema de Heine-Borel

1. Ninguno de ellos.

3. Todos los subconjuntos de M son acotados, de modo que A compacto $\Leftrightarrow$ (A cerrado y acotado) $\Leftrightarrow$ A cerrado.

5. No.

3.3 Propiedad de los conjuntos encajados

1. $\bigcap_k F_k = \{\sqrt{2}\} \neq \varnothing$.

3. Si $F_k = \{x_l \mid l \geq k\}$, entonces $\bigcap_k F_k = \varnothing$. Ninguno de los conjuntos F_k es compacto.

3.4 Conjuntos conexos por arcos

1. **a.** No es conexo por arcos. Cualquier arco entre dos racionales debe contener un irracional.

   b. Conexo por arcos.

   c. Conexo por arcos.

   d. No es conexo por arcos. Si agregamos el punto (1, 0), sería conexo por arcos.

3. No.

3.5 Conjuntos conexos

1. No. $]-1/2, 3/2[$ y $]2, 7/2[$ son conjuntos abiertos disjuntos cuya unión contiene a A.

3. Si $A \subset \mathbb{R}^2$, sea $A^* = \{(x, y, 0) \in \mathbb{R}^3 \mid (x, y) \in A\}$. Si $\gamma(t) = (x(t), y(t))$ es un arco en A, entonces $\gamma^*(t) = (x(t), y(t), 0)$ es un arco en A^*. Si U^* y V^* separan A^*, entonces $U = \{(x, y) \in \mathbb{R}^2 \mid (x, y, 0) \in U^*\}$ y $V = \{(x, y) \in \mathbb{R}^2 \mid (x, y, 0) \in V^*\}$ separa A.

Ejercicios del capítulo 3 (al final del capítulo)

1. **a.** Conexo, no compacto.

   b. Conexo y compacto.

   c. Conexo y compacto.

   d. Ni conexo ni compacto.

   e. Compacto, pero no conexo, si contiene más de un punto.

f. $n = 1$, compacto y no conexo. $n \geq 2$, compacto y conexo.

g. Compacto y conexo.

h. Compacto, no necesariamente conexo.

i. Ni compacto ni conexo.

j. Compacto, no necesariamente conexo.

3. $\text{cl}(A)$ está acotado pues A lo está (¿por qué?), y es cerrado. Por lo tanto, es compacto. Todo subconjunto infinito, como A, debe tener un punto de acumulación (véase la sección 3.1, ejercicio 1).

5. **a.** $U_k = \{x \in \mathbb{R}^2 \mid \|x\| < k/(k+1)\}$, $k = 1, 2, 3, \ldots$.

b. $U_k =]k - (1/3), k + (1/3)[$, $k = \ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$.

7. Comiéncese con $\text{cl}(A_k) = \{x\} \cup \{x_k, x_{k+1}, \ldots\}$. Recuérdese demostrar que si $y \neq x$, entonces $y \notin \text{cl}(A_k)$ para k suficientemente grande (ninguno de tales x_k puede estar cerca de y, pues están cerca de x). Proporciónense detalles.

9. **a.** Falso; $[0, 1]$ es compacto, pero $\mathbb{R} \setminus [0, 1]$ no es conexo. En $\mathbb{R}^n$, $A = \{x \in \mathbb{R}^n \mid 1 \leq \|x\| \leq 2\}$ es compacto, pero $\mathbb{R}^n \setminus A$ no es conexo.

b. Falso; mismos ejemplos que en **a.**

c. Falso; $]a, b]$ es conexo pero no es abierto ni cerrado.

d. Falso para $n = 1$, verdadero para $n \geq 2$. ($\mathbb{R}^n \setminus A$ es conexo por arcos si $n \geq 2$.)

11. **a.** Si U y V separan B y $x \in B \cap U$, entonces $A \cap U \neq \varnothing$ (¿por qué?). Análogamente, $A \cap V \neq \varnothing$. Así, A sería disconexo.

b. Establézcase el siguiente lema: si B es conexo, $B \subset C$, y C es separado por U y V, entonces $B \subset U$ o $B \subset V$.

13. Sea $x_n \in F_n$. Muéstrese que $\langle x_n \rangle_1^\infty$ es una sucesión de Cauchy. Su límite debe estar en $\text{cl}(F_n)$ para todo n (¿por qué?). No pueden existir dos de tales puntos, pues $\text{diam}(F_n) \to 0$.

15. **a.** Si $\gamma(t)$ y $\mu(t)$ son arcos de x a a en K_1 y de y a b en K_2, entonces $\varphi(t) = (\gamma(t), \mu(t))$ es un arco de (x, y) a (a, b) en $K_1 \times K_2$.

17. Como en el ejemplo resuelto **1.2** del capítulo 1, encuéntrese $z_k \in K$ tal que $d(x, z_k) \to d(z, K) = \inf\{d(x, z) \mid z \in K\}$. Para k grande, todos estos puntos están en la bola cerrada de radio $1 + d(z, K)$ centrada en x. Úsese la compacidad para obtener una subsucesión convergente a algún z. Entonces $z \in K$ (¿por qué?) y $d(z_{n(j)}, x) \to d(z, x)$ (¿por qué?). Así, $d(x, z) = d(x, K)$ (¿por qué?). Esto no sirve para conjuntos abiertos. La demostración no sirve a menos que las bolas cerradas sean compactas.

19. $\mathrm{cl}(V_{n+1}) \subset V_n \subset \mathrm{cl}(V_n)$. Úsese la propiedad de los conjuntos encajados.

21. Analícense con cuidado las definiciones para el enunciado general. Sabemos que $\mathbb{R}^n$ es conexo por arcos, por lo cual es conexo.

23. $\mathbb{Q} \subset]-\infty, \sqrt{2}[\cup]\sqrt{2}, \infty[$; ambos intervalos son abiertos y son disjuntos. Separan $\mathbb{Q}$. Análogamente $\mathbb{R}\backslash\mathbb{Q} \subset]-\infty, 0[\cup]0, \infty[$ separan $\mathbb{R}\backslash\mathbb{Q}$.

25. La sucesión sen n para $n = 1, 2, 3, \ldots$ está contenida en el conjunto compacto $[-1, 1]$ y, por lo tanto, tiene una subsucesión convergente sen (n_k).

27. Si $x \in \mathrm{cl}(A)$, existe una sucesión $\langle x_n \rangle_1^\infty$ en A convergente a x. Si se cumple la condición, entonces $x \in A$ (¿por qué?). Así, $\mathrm{cl}(A) \subset A$ y A es cerrado. Para el recíproco, si A es cerrado y acotado, el límite inferior y el límite superior de una sucesión en A son los límites de subsucesiones, por lo que están en A.

29. A es compacto y conexo.

31. Sugerencia: un conjunto en $\mathbb{R}^n$ es compacto sii todo recubrimiento abierto numerable contiene un subrecubrimiento finito. Para obtener esto, obsérvese que todo conjunto abierto es una unión de bolas con centros y radios racionales (¿por qué?).

33. Si $A_1, A_2, A_3, \ldots$ son conjuntos cerrados nunca densos, defínase $B_k = \mathbb{R}^n \backslash A_k$ para ver que la afirmación es equivalente a

Teorema *Si $B_1, B_2, \ldots$ son subconjuntos abiertos densos de $\mathbb{R}^n$, entonces $B = \bigcap_k B_k$ es denso en $\mathbb{R}^n$.*

Sean $x \in \mathbb{R}^n$ y $\varepsilon > 0$. Definimos inductivamente una sucesión encajada de discos cerrados $B(x_k, r_k) = \{y \mid \|x_k - y\| \leq r_k\}$ como

$$D_1 = B(x_1, r_1), \text{ donde } x_1 \in B_1,\ \|x - x_1\| < \varepsilon/3,\quad r_1 < \varepsilon/3 \quad \text{y} \quad B(x_1, r_1) \subset B_1;$$

$$D_2 = B(x_2, r_2), \text{ donde } x_2 \in B_2,\ \|x_1 - x_2\| < r_1/3,\quad r_2 < r_1/3 \quad \text{y} \quad B(x_2, r_2) \subset B_2;$$

etcétera.

Verifíquese que $D_k \subset B_k$ para cada k y que $D(x, \varepsilon) \supseteq \mathrm{cl}(D_1) \supseteq D_1 \supseteq \mathrm{cl}(D_2) \supseteq D_2 \supseteq \cdots$. La existencia de $y \in \bigcap_k \mathrm{cl}(D_k)$ muestra que B es denso en $\mathbb{R}^n$ (¿por qué?).

35. **a.** Todo a. Si $a = 0$ o 1, la sucesión es constante.

b. $0 \leq a \leq 1$.

c. $0 \leq a \leq 1$.

37. **a.** Para cada $x \in A$ existe δ_x tal que $D(x, \delta_x) \subset M \setminus B$; aplíquese la compacidad al recubrimiento $\{D(x, \delta_x/2) \mid x \in A\}$.

b. No; sea A el eje Y y B la gráfica de $y = 1/x$.

Capítulo 4: Transformaciones continuas

4.1 Continuidad

1. **a.** Para $x_0 \in \mathbb{R}$ y $\varepsilon > 0$, inténtese con $\delta = \text{mín}(1, \varepsilon/(1 + 2|x_0|))$.

3. $A = f^{-1}([0, 1])$. Este conjunto es cerrado, pues f es continua y $[0, 1]$ es cerrado.

5. **a.** $f(x) = 1$, U = cualquier conjunto abierto.

b. $f(x) = \begin{cases} 0, & \text{si } x \leq 0. \\ x, & \text{si } 0 < x < 1; \\ 1, & \text{si } x \geq 0. \end{cases}$ $U =]-1, 2[$; $f(U) = [0, 1]$ es cerrado.

4.2 Imágenes de conjuntos compactos y conexos

1. **a.** Cerrado, no necesariamente compacto ni conexo.

b. Abierto, no necesariamente compacto ni conexo.

c. Conexo, no necesariamente compacto, ni abierto ni cerrado.

d. Compacto, cerrado y conexo; no necesariamente abierto.

3. No es posible si B_n es cerrado y acotado, pues entonces es compacto.

5. Sí.

4.3 Operaciones con transformaciones continuas

1. **a.** En todo punto.

b. f es continua en $\mathbb{R}\setminus\{-1, 1\}$.

c. En todo punto.

3. Úsese el hecho de que $\{0.56\}$ es cerrado y sen x es continua. A no es compacto.

5. $f = g \circ h$, donde $g(x) = \sqrt{x}$ y $h(x) = x^2 + 1$. f es continua, pues g y h lo son.

4.4 La acotación de las transformaciones continuas sobre conjuntos compactos

1. $f(x) = x/(1 + |x|)$.

3. $M = f^{-1}(\{\sup(f(x))\})$ ¿Por qué existe? ¿Por qué es compacto?

5. $\sup(f(]0, \infty[)) = 1$ no se alcanza en $]0, \infty[$. Extiéndase mediante $f(0) = 1$ (¿de forma continua?) para obtenerlo en $[0, \infty[$.

4.5 Teorema de los valores intermedios

1. Los polinomios cuadráticos no tienen por qué ser negativos en ningún punto, así que el método falla. El método funciona para polinomios de quinto grado y, en general, para todos los polinomios de grado impar.

3. Aplíquese el teorema de los valores intermedios a $g(x) = f(x) - x$.

5. $f([0, 1])$ sería compacto, y $]0, 1[$ no es compacto.

4.6 Continuidad uniforme

1. $|(1/x) - (1/y)| = |(x - y)/xy| \leq |x - y|/a^2$. Tómese $\delta = a^2\varepsilon$.

3. No. Considérese $f(x) = \text{sen}(x^2)$.

5. $\delta = \varepsilon$ funciona en todo punto.

7. **a.** f es continua en un dominio compacto.

4.7 Derivación de funciones de una variable

1. $f(x) = \sum_{i=1}^{n} a_i |x - x_i|$; $a_i \neq 0$.

3. f alcanza su máximo y su mínimo (¿por qué?). Si ambos se alcanzan en los extremos, entonces f es constante (¿por qué?).

5. Considérese la función $f(x)/x$.

4.8 Integración de funciones de una variable

1. Las sumas inferiores y superiores son todas iguales a $b - a$.

3. Esto es una consecuencia del hecho de que para cualquier $S \subset \mathbb{R}$ tenemos $\sup(S) \geq \inf(S)$, cuando se aplica al conjunto $S = \{f(x) \mid x \in [x_i, x_{i+1}]\}$.

5. $(1/6)(e^{75} - 1)$.

7. $L(f, P) = 0$ para toda partición de $[0, 1]$ (¿por qué?). Úsese ahora una partición cuyo primer intervalo sea $[x_0, x_1] = [0, \sqrt{2}/n]$ y los demás tengan longitud no mayor que $1/n^2$. Muéstrese que $U(f, P) \leq (\sqrt{2}/n) + (n - 1)/n^2$.

Ejercicios del capítulo 4 (al final del capítulo)

1. **a.** $|(1/x^2) - (1/x_0)^2| \leq |x - x_0| (|x| + |x_0|)/x^2x_0^2$. Si $\varepsilon > 0$, inténtese con $\delta = \text{mín}(x_0/2, \varepsilon x_0^3/10)$.

b. Si $\varepsilon > 0$, sea δ cualquier valor mayor que 0.

c. Sí.

3. **a.** No. Sean $f(x) = \text{sen } x$ y $k = \{1\}$.

b. f es continua en todo $\mathbb{R}^n$, por lo que f es continua en $\text{cl}(B)$, que es compacto. $f(\text{cl}(B))$ es compacto y, por tanto, acotado. Así, como $f(B) \subset f(\text{cl}(B))$, $f(B)$ también está acotado.

5. Si f es continua, la condición se cumple por **4.1.1** y **4.1.2**. Recíprocamente, si la condición se cumple para todo $\varepsilon > 0$, entonces se cumple para $\varepsilon/2$, y $\|f(x) - f(x_0)\| \leq \varepsilon/2 < \varepsilon$. No podemos reemplazar " > 0" por " $= 0$".

7. Supóngase que C es un subconjunto cerrado de B. Entonces B es compacto (¿por qué?) y $((f^{-1})^{-1})(C) = f(C)$ (¿por qué?). Así, $f(C)$ es cerrado (¿por qué?). Por lo tanto, f^{-1} es continua (¿por qué?). Para ver un contraejemplo con $n = 2$, considérese $f: [0, 2\pi[\to \mathbb{R}^2$ dada por $f(t) = (\text{sen } t, \cos t)$.

9. Extiéndase f haciendo $f(b) = g(b)$. Si F es cerrado en $\mathbb{R}^m$, muéstrese que $h^{-1}(F) = f^{-1}(F) \cap f^{-1}(F)$ y, que, por lo tanto, es cerrado.

11. **a.** Si $(c, f(c))$ y $(d, f(d))$ están en la gráfica, hágase $\gamma(t) = (t, f(t))$.

13. Si $\inf(f(V))$ o $\sup(f(V))$ estuvieran en $f(V)$, entonces $f(V)$ no podría ser abierto, pues no podría contener ningún intervalo abierto en torno a ninguno de estos puntos.

15. Para un ejemplo de la desigualdad, tómense $f_1(x) = x$ y $f_2(x) = 1 - x$ en $[0, 1]$.

17. Sea $f(x, y) = xy$. Muéstrese que f no es uniformemente continua, mostrando que $|f(a, a) - f(b, b)| = (|a + b|/\sqrt{2}) \|(a, a) - (b, b)\|$.

19. $A = \{(x, y) \in \mathbb{R}^2 \mid xy = 1\}$. $f(A) = \{x \in \mathbb{R} \mid x \neq 0\}$.

21. **a, b** y **c.** Sí. **d.** No.

23. Sugerencia: para mostrar que es suprayectiva, supóngase que $y_1 \in X \setminus f(X)$ y considérese la sucesión $y_2 = f(y_1)$, $y_3 = f(y_2)$, ...

25. $f(x) = \text{sen}(1/x)$ es un contraejemplo con f' no acotada.

27. 81/64.

29. Divida entre $x - y$ y haga tender y a x para mostrar que $f'(x) = 0$.

31. $f(1) = e$.

33. Si A es relativamente compacto, entonces cl(A) es compacto. Las subsucesiones convergentes existen por el teorema de Bolzano-Weierstrass. Si toda sucesión en A tiene una subsucesión convergente en $\mathbb{R}^n$, entonces comiéncese con una sucesión en cl(A), obténgase una sucesión cercana en A, considérese una subsucesión convergente en $\mathbb{R}^n$ y muéstrese que la subsucesión correspondiente de la sucesión original también converge necesariamente a un punto en cl(A).

35. $A = f(\{x \in \mathbb{R} \mid x > 0\})$ no es vacío y está acotado superiormente por $f(0)$. Muéstrese que $\lim_{x \to 0^+} f(x) = \inf(A)$.

37. Supóngase que $f'(a) < 0 < f'(b)$. Como f es continua en $[a, b]$ (¿por qué?), tiene un mínimo en algún x_0 en $[a, b]$ (¿por qué?). $x_0 \neq a$ pues $f(x) < f(a)$ para x ligeramente mayor que a (¿por qué?). $x_0 \neq b$ pues $f(x) < f(b)$ para x ligeramente menor que b (¿por qué?). Así, $a < x_0 < b$, de modo que $f'(x_0) = 0$ (¿por qué?). El caso $f'(a) > 0 > f'(b)$ es análogo.

39. Considérese la segunda derivada de $x^2 f(x)$.

41. Ambas son bastante inmediatas aplicando la inducción matemática. Para la primera, puede procederse partiendo del paso n al paso $n + 1$, sumando $n + 1$ a ambos miembros. Para la segunda, súmese $(n + 1)^2$.

43. Úsese la regla de la cadena y el teorema fundamental del cálculo. Si $F(x) = \int_0^{g(x)} f(y)\, dy$ y g es derivable y f continua, entonces $F'(x) = f(g(x))g'(x)$.

45. Sean $m = \inf(g[a, b]))$ y $M = \sup(g[a, b]))$. Entonces

$$m \int_a^b f(x)\, dx \leq \int_a^b f(x)g(x)\, dx \leq M \int_a^b f(x)\, dx.$$

(¿Por qué?). Como $t \int_a^b f(x)\, dx$ depende continuamente de t, el teorema de los valores intermedios implica que existe t_0 en $[m, M]$ tal que

$$\int_a^b f(x)g(x)\, dx = t_0 \int_a^b f(x)\, dx.$$

Aplíquese ahora el teorema a g para obtener x_0 tal que $g(x_0) = t_0$ (proporciónense detalles).

Capítulo 5: Convergencia uniforme

5.1 Convergencias uniforme y puntual

1. Sí.

3. Sí.

5. Muéstrese que $f_k(x)$ es la k-ésima suma parcial; entonces $|f(x) - f_k(x)| \leq \sum_{n=k+1}^{\infty} (1/n^2)$. Úsese esto para mostrar la convergencia uniforme.

5.2 Criterio M de Weierstrass

1. **a.** La convergencia es puntual pero no uniforme (¿por qué y a qué converge?).

b. La convergencia es uniforme (¿por qué y a qué converge?).

3. $f(x)$ es el límite uniforme de $f_k(x) = \sum_{n=1}^{k} (x^n/n^2)$ en [0, 1] (¿por qué?). Cada f_k es continua, por lo que f también lo es (¿por qué?).

5. Úsese el criterio M de Weierstrass con $M_k = |a_k|$.

5.3 Integración y derivación de series

1. El límite es $f(x) = 1/x$ para $x > 0$, y $f(0) = 0$. Esta función no es continua. La convergencia no puede ser uniforme (¿por qué no?).

3. Muéstrese que $|f_n(x)| \leq f_n(n/(n+1)) \leq \sqrt{n}/(n + 1)$. Con esto muéstrese que $f_n \to 0$ uniformemente en [0, 1]. Las derivadas convergen a 0 puntual pero no uniformemente.

5. Justifíquese la derivación término a término de la serie del seno, analizando la convergencia uniforme de la serie adecuada en $[-R, R]$ para cada $R > 0$.

5.4 Las funciones elementales

1. Justifique la integración término a término analizando la convergencia uniforme en $[0, r]$ para $r > x$.

3. Sea $h(x) = f(ax) - f(a) - f(x)$. Muéstrese que $h(1) = 0$ y $h'(x) = 0$ para todo $x > 0$. Obténgase el resultado de esto.

5. Úsese **4.7.15.**

7. Para los enteros positivos p y q, tenemos $(g(1/q))^q = g((1/q) + \cdots + (1/q))$ (q términos) $= g(1) = E$. Así, $g(1/q) = E^{1/q}$. Entonces $g(p/q) = g((1/q) + \cdots + (1/q))$ (términos p) $= g(1/q)^p = E^{p/q}$.

5.5 El espacio de las funciones continuas

1. No; $\text{int}(B) = \{f \in C_b(\mathbb{R}, \mathbb{R}) \mid \exists\delta > 0\}$ con $f(x) > \delta$ para todo $x\}$.

3. Se puede aplicar el ejemplo **5.5.6** para obtener el criterio M. Si $f_n = \sum_{k=1}^{n-1} g_k$, entonces $\|f_{n+1} - f_n\| = \|g_n\| \le r_n = M_n$.

5. No es cerrado, a menos que la función límite esté en el conjunto.

5.6 Teorema de Arzela-Ascoli

1. Si $|f_n'(x)| < M$ para todo x y $f_n(0) = 0$ para todo n, entonces $|f_n(x)| \le M_x \le M$ para todo $x \in]0, 1[$ (¿por qué?).

3. **a.** Muéstrese que el complementario es cerrado usando **5.6.4.**

b. Un límite uniforme de funciones continuas es continuo.

5. Muéstrese que si $|f_n(t)| \le M$ para todo t en $[a, b]$, entonces $|F_n(x)| \le M(b - a)$ para todo n y todo x en $[a, b]$. Así, $\langle F_n\rangle_1^\infty$ está uniformemente acotada. Además, $|F_n'(x)| \le M$. (¿Por qué?) Úsese el método del ejemplo 1 para obtener una subsucesión uniformemente convergente.

5.7 Principio de la aplicación contractiva y sus aplicaciones

1. $|\alpha| < 1$.

3. $r < 1/2$.

5. $f(x) = 1 + \int_0^x 3sf(s)\,ds$. Sea $T(g)(x) = 1 + \int_0^x 3sg(s)\,ds$ y calcúlese $T(0)(x)$, $T^2(0)(x)$.

7. $x(t) = 1/(1 - t)$. Partiendo de $t = 0$, esta función explota cuando $t \to 1$.

5.8 Teorema de Stone-Weierstrass

1. Por **5.8.4**, los polinomios son densos en $C([0, 2\pi], \mathbb{R})$. Como sen x es continua, podemos tomar $\varepsilon = 1/100$ para obtener el polinomio p pedido.

3. La respuesta a la segunda parte es sí.

5. Sí.

5.9 Criterios de Abel y Dirichlet

1. $\sum_1^\infty (x^n/n!)e^{-nx}$ converge uniformemente por el criterio M, con $M_n = 1/n!$ (obsérvese que $\sum_{n=1}^\infty (1/n!) = e$).

3. Úsese el criterio de Dirichlet para obtener la convergencia uniforme.

5. Diverge en $x = 1$, pero converge uniformemente en $[0, a]$ para $0 < a < 1$.

5.10 Series de potencias y sumabilidad Cesaro y Abel

1. $R = 1, R = 0$.

3. $S_n = 1, 1, 0, 1, 1, 0, \ldots$ y $\sigma_n \to 2/3$.

5. Desarróllese $1/(1 + t^2)$ como una serie geométrica para $|t| < 1$. Después, intégrese término a término para obtener un desarrollo del arco tangente. Justifíquese el uso del teorema de Abel en $x = 1$.

Ejercicios del capítulo 5 (al final del capítulo)

1. **a.** Para $\varepsilon > 0$, tómese K tal que $k \geq K \Rightarrow m_k < \varepsilon$. Entonces $k \geq K \Rightarrow \|f_k(x) - f(x)\| < \varepsilon$ para todo $x \in A$.

3. **a.** No converge en ningún punto.

b. Converge uniformemente en $\mathbb{R}$. El límite es continuo.

c. Converge uniformemente en $\mathbb{R}$ (criterio de Dirichlet). El límite es continuo.

d. Converge en $]0, 1[$ a $x/(1 - x)$, pero no uniformemente.

5. La desigualdad fundamental es

$$|f_k(x)g_k(x) - f(x)g(x)| \leq |f_k(x)|\,|g_k(x) - g(x)| + |f_k(x) - f(x)|\,|g(x)|\,.$$

Considérese $f_k(x) = 1/x$ y $g_k(x) = 1/k$ en $]0, 1[$. $f_k g_k \to 0$ no uniformemente. No.

7. Si $x \in A$, entonces $|f(x)| \leq \|f\|$ y $|g(x)| \leq \|g\|$, por lo que $|f(x)g(x)| \leq \|f\| \cdot \|g\|$.

9. $g(x) = \sum_n g_n(x)$ es continua en A (¿por qué?). Así, $x_k \to x_0 \Rightarrow g(x_k) \to g(x_0)$.

11. **a.** No. Sin la completitud podría no haber punto fijo.

b. No; considérese $f(x) = x + 1/x$ en $[2, \infty[$. Con la compacidad, la respuesta es sí. Considérese $g(x) = d(f(x), x)$. Muéstrese que g debe alcanzar un mínimo en X y que el mínimo debe ser 0.

13. Muéstrese que $|f_k(x) - f(x)| \leq |f_k(c) - f(c)| + \int_c^x |f_k'(t) - g(t)|\, dt$ para $c \in]a, b[$ (inviértase el orden de los límites en la integral si $x < c$). Úsese la convergencia puntual de f_k a f en c y la convergencia uniforme de f_k' a g para hacer la diferencia uniformemente pequeña.

15. Sin la convergencia absoluta, considérese $(1/2) + (1/4) + (1/6) + \cdots$ como una subsuma de $-1 + (1/2) - (1/3) + (1/4) - + \cdots$.

17. Sea $\langle p_n \rangle_1^\infty$ la sucesión de términos positivos y $\langle q_n \rangle_1^\infty$ los términos negativos. Muéstrese que $\sum p_n = +\infty$ y $\sum q_n = -\infty$. Selecciónense términos positivos para el reordenamiento hasta que la suma sea mayor que x y después los términos negativos, hasta que sea menor que x, etcétera, cambiando siempre que la suma cruce x. ¿Por qué puede hacerse esto una y otra vez? ¿Por qué converge el resultado a x?

19. Úsese el criterio M para obtener la convergencia uniforme y la continuidad de la suma en cada intervalo acotado.

21. **b.** Considérese $f_n(x) = x^n$, $\varepsilon = 1/2$ y $x = 1$.

23. No.

25. Úsese el teorema de los valores intermedios (si $f(0) < f(1)$, muéstrese que f es creciente; si $f(0) > f(1)$, muéstrese que f es decreciente) para mostrar que si $x < y < z$ y $f(x) < f(z) < f(y)$, entonces f no es inyectiva.

27. Úsese el método del ejercicio 4, capítulo 4, o el propio ejercicio, apartados **b** y **c**.

29. **a.** f es uniformemente continua en $[-1, 1]$ (¿por qué?), por lo que también lo es en $]-1, 1[$.

b. f es uniformemente continua en $[0, 1]$ y en $[1, \infty]$ (¿por qué?), por lo que también lo es en $[0, \infty]$ (¿por qué?).

c. Sí; f' está acotada.

31. Selecciónese M tal que $|a_k - a| < \varepsilon/2$ para $k \geq M$. Después, elíjase $N > M$ de modo que $|(a_1 - a) + \cdots + (a_M - a)|/N < \varepsilon/2$. Si $n > M$, entonces

$$\begin{aligned} |b_n - a| &= |(a_1 + \cdots + a_n - na)/n| \\ &\leq \frac{|(a_1) + \cdots + (a_m - a)|}{n} + \frac{|a_{M+1} - a|}{n} + \cdots + \frac{|a_n - a|}{n} \\ &\leq \frac{|(a_1) + \cdots + (a_m - a)|}{n} + \frac{n\varepsilon/2}{n} < \varepsilon \end{aligned}$$

33. **a.** Sí.

b. No.

35. Úsese el teorema de los valores intermedios para mostrar que f es estrictamente

monótona. Así, $f([a, b])$ es un intervalo cerrado $[\alpha, \beta]$. La continuidad de f^{-1} en $x_0 \in [\alpha, \beta]$ se puede reducir ahora al caso compacto (capítulo 4, ejercicio 7).

37. Úsese la conexión o el teorema de los valores intermedios.

39. Úsese la continuidad uniforme para elegir subintervalos adecuadamente pequeños. Hágase que g sea constante en estos subintervalos.

41. Úsese **5.8.4** con $\varepsilon = 1/10$.

43. **a.** Para la clausura, verifíquese que el límite de funciones pares continuas es par.

b. Úsese el teorema de Stone-Weierstrass para deducir que los polinomios pares son densos en $C([0, 1])$. Para $f \in C_e([-1, 1])$, aproxímese f mediante un polinomio par p en $[0, 1]$ y después muéstrese que p también aproxima f en $[-1, 0]$. Como $C_e([-1, 1])$ es un subconjunto cerrado propio de $C([-1, 1])$, los polinomios pares no pueden ser densos en $C([-1, 1])$.

45. **a.** Muéstrese primero que f es uniformemente continua en K usando la equicontinuidad. Combínese esto con la compacidad de K y úsese de nuevo la equicontinuidad para obtener la convergencia uniforme.

b. Para ver que la convergencia no es uniforme, obsérvese que $f_n(1/n) = 1$ y que no tiende a 0. La sucesión $\langle f_n \rangle_1^\infty$ no puede ser equicontinua.

47. Esto muestra que la convergencia del subconjunto numerable denso que hemos obtenido implica la convergencia en todo punto.

49. Si $\langle f_n \rangle_1^\infty$ es una sucesión uniformemente de Cauchy, entonces existe una función f a la que f_n converge uniformemente, por **5.2.1**. f está acotada, pues si $|f(x) - f_n(x)| \leq 1$ para todo x, entonces $|f(x)| \leq |f_n(x)| + 1 \leq \|f_n\| + 1$.

51. Podemos sumar las filas y después las sumas de las filas, o bien las columnas y después las sumas de las columnas.

53. **a.** 0.

b. 2.

c. 0.

55. Véase §**5.5**, ejercicio 5.

57. Como $|f_n| \leq g$ para cada n, $|f| \leq g$. Como $\int_0^\infty g$ existe, habrá un $R > 0$ tal que $\left|\int_R^\infty g\right| < \varepsilon/4$. En $[0, R]$, $f_n \to f$ uniformemente, de modo que $\int_0^R f_n \to \int_0^R f$. Elíjase n suficientemente grande como para que $\left|\int_0^R f_n - \int_0^R f\right| < \varepsilon/2$. Entonces $\left|\int_0^\infty f_n - \int_0^\infty f\right| \leq \left|\int_0^R f_n - \int_0^R f\right| + \left|\int_R^\infty f_n\right| + \left|\int_R^\infty f_n\right| \leq \left|\int_0^R f_n - \int_0^R f\right| + 2\left|\int_R^\infty g\right| < \varepsilon$.

59. **c.** Por el apartado **b**, tenemos

$$\begin{aligned}\sum(a_k+b_k)^p &= \sum(a_k+b_k)^{p-1}a_k+\sum(a_k+b_k)^{p-1}b_k\\ &\le \left(\sum(a_k+b_k)^{q(p-1)}\right)^{1/q}\left(\sum a_k^p\right)^{1/p}\\ &\quad+\left(\sum(a_k+b_k)^{q(p-1)}\right)^{1/q}\left(\sum b_k^p\right)^{1/p}\end{aligned}$$

Pero $q(p-1)=p$. Así, $(\sum(a_k+b_k)^p)^{1/p}=(\sum(a_k+b_k)^p)^{1-(1/q)}\le(\sum a_k^p)^{1/p}+(\sum b_k^p)^{1/p}$ como se deseaba.

61. Sabemos que $|x-b|=\rho<R$ y que $\sum|a_k|\rho^k$ converge. En consecuencia, $\sum|a_k(x-b)^k|$ converge por comparación.

63. Sea $u_k=(\alpha(\alpha-1)\cdot\ldots\cdot(\alpha-k+1)x^k)/(k!)$. Si $k>|\alpha|$, entonces $|u_{k+1}/u_k|=|x|(k-\alpha)/(k+1)$. Esto tiende a $|x|$ cuando $k\to\infty$. El criterio del cociente muestra que el radio de convergencia es 1.

65. Inténtese $(C,1)\int_0^\infty f(x)dx=\lim_{B\to\infty}(1/B)\int_0^B(\int_0^r f(x)dx)\,dr$ si este límite existe.

67. Muéstrese que $|f(t_1)-f(t_2)|\le 2\sum_N^\infty\varepsilon_n$ si $|t_1-t_2|<1/2^N$ y véase el ejercicio 24 del capítulo 4.

Capítulo 6: Transformaciones diferenciables

6.1 Definición de derivada

1. $Df(x)=\operatorname{sen}x+x\cos x$ (es decir, la multiplicación por esta cantidad).

3. Sea $f(x,y)=0$. Muéstrese que $Df(x,y)=0$ y $Df(x,y)(h,k)=k$ satisfacen ambas la definición.

5. No.

6.2 Representación matricial

1. $Df(x,y,z)=\begin{pmatrix}4x^3y & x^4 & 0\\ e^z & 0 & xe^z\end{pmatrix}.$

3. Por §**6.1**, ejercicio 4, sabemos que $Dg(0)=0$. Por **6.2.4**, sabemos que $DL(x)=L$ para cada x. Por §**6.1**, ejercicio 2, tenemos que $D(g+L)(0)=Dg(0)+DL(0)=0+L=L$.

5. La definición es la misma, excepto que por lo general pedimos que la transformación lineal en la definición de la derivada sea continua. Esto es automático en $\mathbb{R}^n$ pero no en los espacios de dimensión infinita.

6.3 Continuidad de las transformaciones diferenciables

1. Sí; sí (véase también §**6.1**, ejercicio 4).

3. No.

5. $c'(1) = (6, e, 3)$

6.4 Condiciones para la diferenciabilidad

1. Muéstrese que $\partial f/\partial x$ y $\partial f/\partial y$ son continuas en (0, 0).

3. $z = 0$.

5. Estúdiese el ejemplo 2 de §**6.3.**

6.5 La regla de la cadena

1. $\partial h/\partial x = (\partial f/\partial u)(\partial u/\partial x) + (\partial f/\partial v)(\partial v/\partial x)$;

$\partial h/\partial y = (\partial f/\partial u)(\partial u/\partial y) + (\partial f/\partial v)(\partial v/\partial y) + (\partial f/\partial w)(\partial w/\partial y)$;

$\partial h/\partial z = (\partial f/\partial u)(\partial u/\partial z) + (\partial f/\partial w)(\partial w/\partial z)$;

3. Ambos miembros son iguales a $2xyf'(x^2 + y^2)$.

5. $0 = (d/dx)[F(x, f(x))] = (\partial F/\partial x) + (\partial F/\partial y)(df/dx)$

6.6 Regla del producto y gradientes

1. Sea $g(t) = x_0 + th$, x_0, $h \in \mathbb{R}^n$ y úsese la regla de la cadena.

$$\frac{d}{dt}\left[f(x_0 + th)\right]\Big|_{t=0} = Df(x_0 + 0h)Dg(0) = Df(x_0)h.$$

3. $2x + y = 2$.

5. Sean $(w_1, \ldots, w_n) = w = f\nabla g + g\nabla f$ y $(v_1, \ldots, v_n) = v = \nabla(fg)$. Entonces

$$w_k = f\frac{\partial g}{\partial x_k} + g\frac{\partial f}{\partial x_k} = \frac{\partial}{\partial x_k}(fg) = v_k.$$

Esto es válido para cada k, por lo que $w = v$.

6.7 El teorema del valor medio

1. Si $a < b$, entonces $f(b) - f(a) = f'(x)(b - a)$ para algún x entre a y b. Como $f'(x) > 0$, tenemos $f(b) > f(a)$.

3. **a.** 1.

   b. 1.

5. La afirmación es una consecuencia sencilla de **6.7.1i**. Es fácil encontrar un contraejemplo si el dominio no es conexo. Sea $A = \mathbb{R} \setminus [0, 1]$. Hágase $f(x) = 1$ si $x > 1$ y $f(x) = 0$ si $x < 0$. Entonces f es derivable en A y $f' = 0$ en A, pero f no es constante en A.

6.8 Teorema de Taylor y derivadas de orden superior

1. $\dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{\partial^2 f}{\partial y \partial x} = 2(2x^3y^3 + x^2y^5 + 3y^3x + 2x^2y + y)e^{x^2+y^2x}$.

3. Esta función es derivable en 0, pero f' no es continua en 0.

5. $f(h, k) = 1 + h + (h^2/2) - (k^2/2) + R_2((h, k), 0)$, donde $R_2((h, k), 0) / \|(h, k)\|^2 \to 0$ cuando $(h, k) \to (0, 0)$.

6.9 Máximos y mínimos

1. Arguméntese como en **6.9.5** o aplíquese **6.9.5** a la matriz $\begin{pmatrix} -a & -b \\ -b & -d \end{pmatrix}$.

3. Mínimo local.

5. Sean $e_1 = (1, 0, 0)$ y $e_3 = (0, 0, 1)$. Entonces $\langle Ae_1, e_1 \rangle = 1$, pero $\langle Ae_3, e_3 \rangle = -1$. Por lo tanto, A no es semidefinida.

Ejercicios del capítulo 6 (al final del capítulo)

1. Muéstrese que $\alpha Df(x_0) + \beta Dg(x_0)$ satisface la definición de diferenciabilidad de $\alpha f + \beta g$ en x_0 y aplíquese la unicidad (**6.1.2**).

3. Si $f(x)$ es idénticamente nula, entonces $f'(c) = 0$ para todo c. En caso contrario, existe x_0 tal que $f(x_0) \neq 0$. Elíjase B de modo que $|f(x)| < |f(x_0)|/2$ para $x \geq B$. Entonces f debe tener un extremo local en algún punto entre 0 y B. En este punto, la derivada debe anularse.

5. **a.** $(2x \cos(x^2 + y^3), 3y^2 \cos(x^2 + y^3))$.

c. (y, x).

e. $\begin{pmatrix} y\cos(xy) & x\cos(xy) \\ -y\,\text{sen}(xy) & -x\,\text{sen}(xy) \\ 2xy^2 & 2x^2y \end{pmatrix}$

g. (yz, xz, xy).

7. **a.** (3, 6) es un mínimo local y (1, 2) es un punto silla.

c. Los puntos $((2k+1)\pi/4, m\pi, 0)$ son puntos silla. Los puntos $(n\pi/2, (2j+1)\pi/2, 0)$ son puntos silla si $n+j$ es par y mínimos locales si $n+j$ es impar.

9. **a.** Inténtese con $\{(x, y) \in \mathbb{R}^2 \mid -1 \le x \le 1 \text{ y } -2 \le y \le 2\} \cup \{(x, y) \in \mathbb{R}^2 \mid -2 \le x \le 2 \text{ y } -1 \le y \le 1\}$.

b. Si (1) se cumple, la demostración de **6.7.2** muestra que $f(x) = f(x_0)$ para todo x en A. Con (2), sea $\gamma(t)$ el arco de x_0 a x en A. Sea $h(t) = f(\gamma(t))$. Verifíquese que $h'(t) = 0$ para todo t y concluyase que $f(x) = f(x_0)$ para todo x en A.

c. Modifíquese la demostración de que un conjunto abierto conexo es conexo por arcos.

11. Las mismas técnicas funcionan en este caso.

13. **a.**

$$\begin{aligned}
\frac{\partial F}{\partial x} &= 2u\frac{\partial u}{\partial x} + w\frac{\partial v}{\partial x} + v\frac{\partial w}{\partial x} = \text{sen}(2x) + yz \\
\frac{\partial F}{\partial y} &= 2u\frac{\partial u}{\partial y} + w\frac{\partial v}{\partial y} + v\frac{\partial w}{\partial y} = xz + 3y^2z \\
\frac{\partial F}{\partial z} &= 2u\frac{\partial u}{\partial z} + w\frac{\partial v}{\partial z} + v\frac{\partial w}{\partial z} = y^3 + xy \\
DF(x, y, z) &= (yz + \text{sen}(2x), xz + 3y^2z, y^3 + xy)
\end{aligned}$$

15. Si S fuera un subconjunto infinito de [0, 1] tendría un punto de acumulación $c \in [0, 1]$ (¿por qué?). Existiría x_k en S tal que $x_k \to c$. Entonces $f(c) = \lim_{k\to\infty} f(x_k) = 0$ y $f'(c) = \lim_{k\to\infty} [(f(x_k) - f(c))/(x_k - c)] = \lim_{k\to\infty} (0/(x_k - c)) = 0$. De este modo, $f(c) = f'(c) = 0$. Pero esto no ocurre en [0, 1]; por lo tanto, S debe ser finito.

17. Imítese la demostración del teorema **6.9.4.**

19. **a.** $f(x, y) = x^2 + 2xy + y^2 + 0$.

b. $f(x, y) = 1 + x + y + (1/2)(x^2 + 2xy + y^2) + R_2(x, y)$

21. Ésta es una consecuencia directa de los teoremas **6.9.4** y **6.9.5**.

23. $0 + 0 - (1/2)x^2 + 0 - (2/4!)x^4$.

25. Revísese la demostración del teorema **6.4.1** y obsérvese que la continuidad de $\partial f/\partial x_n$ no es necesaria.

27. Considérese $f(x, y) = (x^2y\sqrt{x^2 + y^2})/(x^4 + y^2)$.

29. **a.** Para $x > 0$, $\sum_{n=0}^{\infty} e^{-nx}$ es una serie geométrica convergente. Multiplíquese por x para obtener $f(x) = xe^x/(e^x - 1)$ para $x > 0$.

b. No, $f(0) = 0$, pero la regla de l'Hôpital implica que $\text{lím}_{x\to 0} f(x) = 1$.

c. $[a, \infty[$ con $a > 0$.

d. Sí, en $]0, \infty[$.

31. δ_{x_0} es lineal y continua, por lo que es diferenciable y $D\delta_{x_0}(f) = \delta_{x_0}(f)$ para todo f.

33. Analícese el desarrollo de Taylor de $f(x) = (a + x)^n$ en torno a $x_0 = 0$.

35. Si $f(x_1) = f(x_2) = f(x_3) = 0$ para $x_1 < x_2 < x_3$, entonces el teorema de Rolle implica $x_1 < c_1 < x_2 < c_2 < x_3$ con $f'(c_1) = f'(c_2) = 0$. Úsese de nuevo el teorema de Rolle para obtener $c_1 < c < c_2$ con $f''(c) = 0$.

37. En un máximo local, $\partial^2 f/\partial x^2 \geq 0$ y $\partial^2 f/\partial y^2 \geq 0$, por lo que ambas deben anularse. Analícese ahora el determinante hessiano.

39. **a.** Punto silla.

b. Un punto silla si $D = C^2 + 4AB > 0$; un mínimo local si $D < 0$ y $A > 0$; un máximo local si $D < 0$ y $A < 0$.

41. $f(x) = x^{5/3}$.

Capítulo 7: Teoremas de la función inversa e implícita y temas relacionados

7.1 Teorema de la función inversa

1. $\dfrac{\partial(u, v)}{\partial(x, y)} = 4x^2 + 4y^2 \neq 0$ excepto en $(x, y) = (0, 0)$.

3. $f'(0) = \text{lím}_{x\to 0}(1 + 2x \operatorname{sen}(1/x)) = 1 \neq 0$. Pero $\text{lím}_{x\to 0} f'(x) = \text{lím}_{x\to 0}(1 - 2\cos(1/x))$ no existe. Por lo tanto, f' no es continua en 0. f no es localmente invertible cerca de 0, pues no es inyectiva en cualquier vecindad de 0 (¿por qué no?)

5. $\left.\dfrac{\partial(u, v, w)}{\partial(x, y, z)}\right|_{(0,0,0)} = 1$. Así, el sistema es invertible cerca de (0, 0, 0).

7.2 Teorema de la función implicita

1. $y = (-1/2) \pm (\sqrt{3}/2)\sqrt{(-1 - 4x)}$; podemos despejar y de manera única en términos de x cerca de todos los puntos de la parábola, excepto (–1/4, –1/2). Existen dichos puntos si $x < -1/4$.

3. $F_1 = 3x + 2y + z^2 + u + v^2$; $F_2 = 4x + 3y + z + u^2 + v + w + 2$; $F_3 = x + z + w + u^2 + 2$, así que

$$\left.\frac{\partial(F_1, F_2, F_3)}{\partial(u, v, w)}\right|_{(0,0,0,0,0,-2)} = 1 \neq 0.$$

de modo que u, v, w se pueden expresar en términos de x, y, z para (x, y, z) en alguna vecindad de (0, 0, 0).

5. $F_1 = y + x + uv$; $F_2 = uxy + v$. Así,

$$\left.\frac{\partial(F_1, F_2)}{\partial(u, v)}\right|_{(0,0,0,0)} = 0.$$

El teorema de la función implícita no garantiza que haya una solución local.

7.3 Teorema de rectificación del dominio

1. $Df(x, y) = (2x, -2y) = 0$ solamente en (0, 0). La función se puede "rectificar" cerca de todo punto, excepto (0, 0).

3. $Df(0, 1) = (0, 2)$. Así, f se puede rectificar cerca de (0, 1). $Df(0, 0) = (0, 0)$, por lo que el teorema no se aplica a este caso. La curva de nivel que pasa por (0, 0) tiene una cúspide ahí y no es suave.

7.4 Más consecuencias del teorema de la función implícita

1. Sí, cerca de (0, 1); no, cerca de (0, 0).

3.
$$Df(0,0,0) = \begin{pmatrix} 1 & 2 & 0 \\ 6 & 12 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

tiene rango 2 y el rango sigue siendo 2 en la bola de radio $\sqrt{2/3}$. Además, $f(0, 0, 0) = (0, 0, 0)$. Así, existen vecindades de $(0, 0, 0)$ y cambios de variable suaves en dichas vecindades tales que $(g \circ f \circ h)(x, y, z) = (x, y, 0)$ cerca de $(0, 0, 0)$.

7.5 Un teorema de existencia para ecuaciones diferenciales ordinarias

1. $f(t, x) = 1 + x^2$. Queremos que $x(t) = \int_0^t (1 + x(s)^2)\, ds$. Sean $x_0(t) = 0$; $x_1(t) = \int_0^t (1 + x_0(s)^2)\, ds = t$; $x_2(t) = \int_0^t (1 + x_1(s)^2)\, ds = t + (1/3)t^3$, y así sucesivamente. La solución exacta es $x(t) = \tan(t)$.

3. $f(x, t) = \sqrt{x}$ no es Lipschitz en ninguna vecindad de $x = 0$.

5. **a.** $e^{tA} = \sum_{n=0}^{\infty} (t^n/n!)A^n$ es absolutamente convergente al igual que su derivada término a término. Así, $(d/dt)(e^{tA}) = \sum_{n=1}^{\infty} (nt^{n-1}/n!)A^n = A \sum_{k=0}^{\infty} (t^k/k!)A^k = Ae^{tA}$.

b. Desplácese el origen del tiempo: $e^{tA} = e^{(t-b)A}e^{bA}$. Obtenemos tiempos $b, 2b, \ldots$.

7.6 Lema de Morse

1. 1.

3. $\det H = 0$. El origen es un punto crítico degenerado. El lema de Morse no es aplicable.

5. **a.** Úsese **7.6.1** y el hecho de que los puntos críticos se "conservan" mediante un cambio de coordenadas.

b. $\{(x, y) \in \mathbb{R}^2 \mid x = 0 \text{ o } y = 0\}$.

7.7 Extremos condicionados y multiplicadores de Lagrange

1. $\sqrt{2/3}(1, -1, 1)$ es un máximo y $\sqrt{2/3}(-1, 1, -1)$ es un mínimo.

3. $(\pm\sqrt{3}, 0)$.

5. Máximo en $(1/\sqrt{70})(9, 4)$, mínimo en $-(1/\sqrt{70})(9, 4)$.

Ejercicios del capítulo 7 (al final del capítulo)

1. $\dfrac{\partial f}{\partial x} = \dfrac{\partial g}{\partial u} + \dfrac{\partial g}{\partial v}\dfrac{\partial h}{\partial x}.$

3. $\dfrac{\partial(u, v)}{\partial(x, y)} = f'(x)$. Así, si $f'(x_0) \neq 0$, el sistema es localmente invertible.

5. **a.** $\dfrac{\partial(x, y)}{\partial(r, \theta)} = r\cos^2\theta + r\,\mathrm{sen}^2\theta = r.$

b. Se puede definir localmente una inversa suave, excepto en torno a $r = 0$ (el origen).

9. Calcúlese $Jf(x, y) = (\partial f_1/\partial x)^2 + (\partial f_1/\partial y)^2$. Así, $Jf = 0 \Leftrightarrow \partial f_1/\partial x = \partial f_1/\partial y = 0$. Esto obliga a que $\partial f_2/\partial y = \partial f_2/\partial x = 0$ por las ecuaciones de Cauchy-Riemann. Así, $Df = 0$.

11. **a.** Úsese **7.2.1.**

b. Úsese **7.4.1.**

15. Derívese la relación $\mathcal{L}^{-1}(A) \circ A =$ identidad con respecto de A.

17. No.

19. Necesitamos que $xy \neq 9u^2v^2$. $\partial u/\partial x = (3v^2 + x)/(xy - 9u^2v^2)$.

23. **a.** C es una unión de rayos que parten del origen (un cono generalizado).

b. Sea $K = \{u \in C \mid \|u\| = 1\} \cup \mathbb{R}^n$. La función continua $\|f\|$ alcanza un máximo en el conjunto compacto K. $\|f(u)\| \leq M$ para todo u en K. Muéstrese que $\|f(v)\| \leq M\ \|v\|$ para todo v en C.

25. Muéstrese que f es una contracción propia de $B(0, r)$ en sí misma y aplíquese el principio de la aplicación contractiva.

29. Consúltese §**5.8** y §**5.9.**

31. $x(t) = \sum_{k=2}^{\infty} = [1/(3 \cdot 5 \cdot \ldots \cdot (2k-3))]t^{2k-3}$; $R \to \infty$.

35. Un valor mínimo igual a 1 en $(\pm 1, 0)$, sin máximos.

35. Un mínimo igual a 0 en $(0, 0, 0)$. Un máximo igual a 4 en $(0, 0, -2)$.

39. Radio $= (500/\pi)^{1/3} \approx 5.42$ cm. Altura $= 2 \times$ radio ≈ 10.84 cm.

Capítulo 8: Integración

8.1 Funciones integrables

1. Sea P una partición en subrectángulos R_j. Si $x_j \in R_j$ son los puntos de evaluación de una suma de Riemann, entonces $L(f, P) = \sum_j \inf(f(R_j))v(R_j) \leq \sum_j f(x_j)v(R_j) \leq \sum_j \sup(f(R_j))v(R_j) = U(f, P)$.

3. Si $\varepsilon > 0$, dividimos $[0, 1]$ como $\{0, (1/2) - (\varepsilon/4), (1/2) + (\varepsilon/4), 1\}$. Entonces $L(f, P) = 0$ y $U(f, P) = \varepsilon/2$. Así, $U(f, P) - L(f, P) = \varepsilon/2 < \varepsilon$. Entonces, f es integrable por la condición de Riemann. Por lo tanto,

$$\int_0^1 f(x)\,dx = \sup\{L(f, P) \mid P \text{ es una partición}\} = 0.$$

5. $11/2$.

8.2 Volumen y conjuntos de medida nula

1. Cada semicircunferencia es la gráfica de una función continua en $[-1, 1]$. Tal gráfica tiene área nula. Recúbrase ésta con rectángulos, usando la continuidad uniforme para dividir $[-1, 1]$ en intervalos lo suficientemente pequeños como para que $f(x)$ cambie no más de $\varepsilon/2$ en cada subintervalo.

3. Si A y $[a, b]\backslash A$ tienen ambos medida nula, entonces también $[a, b]$, por **8.2.4.**

5. No necesariamente. Considérese $A = \mathbb{Q} \cap [0, 1] \subset \mathbb{R}$.

8.3 Teorema de Lebesgue

1. El conjunto acotado $[-1, 1]$ tiene volumen (longitud) y f es continua y acotada en él. Por lo tanto, f es integrable, por **8.3.3.**

3. 1.

5. Sí.

8.4 Propiedades de la integral

1. No necesariamente. Un único punto tiene volumen 0, pero $\mathbb{Q} \cap [0, 1]$ es un conjunto numerable acotado que no tiene volumen.

3. $1_{A\cup B} = 1_A + 1_B - 1_{A\cap B}$ (¿por qué?). Así, $v(A\cup B) = \int 1_{A\cup B} = \int 1_A + \int 1_B - \int 1_{A\cap B} = v(A) + v(B) - v(A\cap B) = v(A) + v(B)$.

8.5 Integrales impropias

1. $0 \leq e^{-x}x^p \leq x^p$. Úsese el criterio de comparación y el ejemplo **2b.**

3. $\alpha < -1$.

8.6 Algunos teoremas de convergencia

1. Considérese el teorema de Dini.

3. 2.

8.7 Introducción a las distribuciones

1. $\delta''(f) = (\delta')'(f) = \delta'(-f') = \delta(-(-f')') = \delta(f'') = f''(0)$.

3. $T_n'(f) = -T_n(f') \to -T(f') = T'(f)$. Esto se cumple para cada f en $\mathcal{D}$ y la convergencia de las distribuciones se define como la convergencia puntual. Así, $T_n' \to T'$.

Ejercicios del capítulo 8 (al final del capítulo)

1. Si h está acotada por M, es integrable y $v(S) = 0$, entonces $\int_S h = 0$, pues $\left|\int_S h\right| \leq \int_S |h| \leq Mv(S) = 0$. Aplíquese esto a f y g. $\int_A f = \int_{A\setminus S} f + \int_S f = \int_{A\setminus S} f = \int_{A\setminus S} g = \int_{A\setminus S} g + \int_S g = \int_A g$.

3. Divídase $[a, b]$ mediante $x_j = a + (j/n)(b - a)$, con $j = 0, 1, 2, \ldots$. Entonces $U(f, P) = [f(x_1) + f(x_2) + \cdots + f(x_n)](b - a)/n$ y $L(f, P) = [f(x_0) + f(x_1) + \cdots + f(x_{n-1})](b - 1)/n$. Así, $U(f, P) - L(f, P) = (f(b) - f(a))(b - a)/n \to 0$ cuando $n \to \infty$. Aplíquese la condición de Riemann.

5. Úsense las cajas $S_n = [-n, n] \times [-n, n] \times [-\varepsilon/n^2 2^{n+4}, \varepsilon/n^2 2^{n+4}]$ para mostrar que todo el plano tiene medida tridimensional nula. Así, cualquier subconjunto cumple la misma propiedad.

7. Úsese la continuidad para mostrar que f debe ser positiva en algún punto, pues es negativa cerca de b y tiene integral 0. Digamos que $f(x_0) > 0$. Por el teorema de los valores intermedios, existe un punto x_1 tal que $a < x_0 < x_1 < b$ y $f(x_1) = 0$. Por el teorema de Rolle, existe un c tal que $a < c < x_1$ y $f'(c) = 0$.

9. **a.** 0.

b. $2b$ sen $b - (b^2 - 2)$ cos $b - 2a$ sen $a + (a^2 - 2)$ cos a.

15. No.

17. $L(f, P_n)$ son crecientes y están acotadas superiormente (¿por qué?). $U(f, P_n)$ son decrecientes y están acotadas inferiormente (¿por qué?). Si los límites son iguales, entonces f será integrable, por la condición de Riemann.

19. Esto no es necesariamente cierto si f no siempre es no negativa.

21. Para $p > 1$, compárese con $\int_1^\infty x^{-p}\,dx$.

23. **a.** En primer lugar, considérese el caso $n = 1$ y úsese la continuidad uniforme.

b. Úsese gráfica$(\varphi) = \bigcup_{n=1}^\infty$ gráfica $(\varphi|[-n, n]|)$.

25. En caso contrario, obténgase $A > 0$ y x_n con $x_{n+1} > x_n + 1$ y $f(x_n) > A$. Úsese la continuidad uniforme para obtener intervalos de longitud 2δ centrados en cada x_n tales que $f(x) > A/2$ para $x_n - \delta \le x \le x_n + \delta$. Dedúzcase de esto que $\int_0^\infty f \ge \sum_{n=1}^N (2\delta)\,A/2 = N\delta A \to \infty$ cuando $N \to \infty$.

27. Si $f(x_0) > 0$, existe $\delta > 0$ tal que $f(x) > f(x_0)/2$ en $D(x_0, \delta) = D$. Así, $\int_D f \ge f(x_0)v(D)/2 > 0$, una contradicción.

29. Los intervalos disjuntos eliminados tienen longitud total $(1/3) + (2/9) + (4/27) + (8/81) + \cdots = 1$.

31. Sea $a = x_0 < x_1 < \cdots < x_n = b$ cualquier partición de $[a, b]$. Por el teorema del valor medio, existen puntos $t_j \in [x_j, x_{j+1}]$ tales que $f(x_{j+1}) - f(x_j) = f'(t_j)(x_{j+1} - x_j)$. Así,

$$f(b) - f(a) = \sum_{j=0}^{n-1}(f(x_{j+1}) - f(x_j)) = \sum_{j=0}^{n-1} f'(t_j)(x_{j+1} - x_j).$$

Ésta es una suma de Riemann para $\int_a^b f'(t)\,dt$; y éstas convergen a dicha integral. Así, $\int_a^b f'(t)\,dt = f(b) - f(a)$.

33. $|f(x)| = 1$ para todo x en $[0, 1]$, por lo que $|f|$ es integrable. Pero $f(x)$ toma los valores ± 1 en cada intervalo. La integral inferior de f es -1 y la integral superior es 1.

35. $(1/n)A_n = \sum_{j=1}^n (1 + (j/n))(1/n)$. Éstas son sumas de Riemann que convergen a $\int_0^1 (1 + x)\,dx = 3/2$.

37. **a.** Por ejemplo: f es continua y φ es C^1.

b. $\sqrt{(1 + x)/(1 - x)}$ + constante.

39. Escríbase la expresión entre corchetes como

$$\frac{1}{n}\sum_{k=1}^{n}\left(\frac{1}{1+(k/n)}\right)$$

y úsense sumas de Riemann.

41. En absoluto; $d(f, g)$ puede ser 0 aunque $f \neq g$.

43. $-\pi \log 2$

47. $1/(\alpha + 1)$

51. Cierto.

53. Muéstrese por inducción que $2^k - k \geq 2^{k-1}$, de modo que $0 < 1/(2^k - k\,\mathrm{sen}(kx)) \leq 1/2^{k-1}$. Úsese entonces el criterio M de Weierstrass.

Capítulo 9: Teorema de Fubini y la fórmula del cambio de variables

9.1 Introducción

1. $(e - 1)/3$

2. $5/6$.

3. $1/2$.

9.2 Teorema de Fubini

3. $\int_1^e \int_{\ln y}^1 (x + y)\, dxdy = (e^2 - 1)/4$.

9.3 Teorema del cambio de variables

3. $\int_0^{1/2} \int_v^{1-v} (2u^2 + 2v^2)\, dudv = 1/3$.

9.4 Coordenadas polares

1. $\pi(e - 1)$.

9.5 Coordenadas esféricas y cilíndricas

1. $\int_D e^{(x^2+y^2+z^2)^{3/2}}\, dxdydz = \int_0^{2\pi} \int_0^{\pi} \int_0^1 e^{r^3} r^2 \operatorname{sen}\varphi\, drd\varphi d\theta = 4\pi(e-1)/3.$

3. $(2\sqrt{2}-1)\pi.$

5. $\pi/4.$

9.7 Operaciones de intercambio de límites

1. **b.** $-1/(1+t)^2.$
   c. $(-1)^n/(1+t)^{n+1}.$

3. $[2\pi t \operatorname{sen}(2\pi t) + \cos(2\pi t) - 1]/t^2.$

5. $(n!)a^{-(n+1)}.$

Ejercicios del capítulo 9 (al final del capítulo)

1. $\pi/3.$

3. **b.** $\pi(1-\cos 1).$
   c. $\infty.$
   d. 2.
   e. 1.
   f. $\pi/3.$
   g. 0.

5. **b.** $(1/3)\pi r_0^2 h.$
   d. $(9\pi)/8.$

7. $(1/4)(1-\cos 1).$

9. Muéstrese que $\tilde{f}\tilde{g}$ es integrable. Entonces

$$\begin{aligned}
\int_{[a,b]\times[c,d]} \tilde{f}(x,y)\tilde{g}(x,y)\, dxdy &= \int_a^b \left(\int_c^d f(x)g(y)\, dy \right) dx \\
&= \int_a^b f(x) \left(\int_c^d g(y)\, dy \right) dx \\
&= \left(\int_a^b f(x)\, dx \right) \left(\int_c^d g(y)\, dy \right).
\end{aligned}$$

11. Si $x \notin \mathbb{Q}$, entonces $1_S(x, y) = 0$. Si $x \in \mathbb{Q}$, entonces $1_S(x, y) \neq 0$ sólo para un número finito de valores de y. Así, $\int_0^1 1_S(x, y)\, dy = 0$ para cada x y $\int_0^1 \left(\int_0^1 1_S\, dy\right) dx = 0$. Pero si P es una partición de $[0, 1] \times [0, 1]$, entonces $L(1_S, P) = 0$ y $U(1_S, P) = 1$ (¿por qué?). Así, 1_S no es integrable (¿por qué no?).

13. Para cada x, $v(C_x) = \int_B 1_{C_x}(y)\, dy$, y

$$v(C) = \int_{A\times B} 1_C = \int_A \left(\int_B 1_{C_x}(y)\, dy\right) dx = 0.$$

Pero $\int_B 1_{C_x}(y)\, dy \geq 0$, de modo que $\int_B 1_{C_x}(y)\, dy = 0$ excepto en un conjunto de medida nula. Si $C = \{(1/2, y) \mid y \in [0, 1]\}$, entonces $v(C) = 0$, pero $v(C_{1/2}) = 1$.

15. Sea $f_n = 1_{A_n}$. Entonces $0 \leq f_n(x) \leq f_{n+1}(x) \leq 1$ y $f_n(x) \to 1_A(x)$ para cada x. Como 1_A es integrable, el teorema de la convergencia monótona implica $v(A_n) = \int f_n \to \int 1_A = v(A)$.

17. Sea $f_M(x) = f(x)$ si $f(x) \leq M$ y 0 si $f(x) > M$. Entonces $\int_0^1 f_M = 0$ pues $f_M \neq 0$ sólo en un número finito de puntos. Entonces

$$\int_0^1 f(x)\, dx = \lim_{M\to\infty} \int_0^1 f_M = 0.$$

19. Justifíquense los pasos del siguiente cálculo.

$$\begin{aligned}
\int_{A\times B} F(x,y)\, dxdy &= \int_{A\times B} f(x)\, dxdy + \int_{A\times B} g(y)\, dxdy \\
&= \int_A \left(\int_B f(x)\, dy\right) dx + \int_B \left(\int_A g(y)\, dx\right) dy \\
&= \int_A \left(f(x) \int_B dy\right) dx + \int_B \left(g(y) \int_A dx\right) dy \\
&= \int_A v(B) f(x)\, dx + \int_B v(A) g(y)\, dy = v(B) \int_A f + v(A) \int_B g.
\end{aligned}$$

21. Úsense las coordenadas esféricas para transformar la integral a $4\pi \int_0^\infty r^{p+2}\, dr$. Esto no converge para ningún p (capítulo 8, ejercicio 15).

23. Como C tiene volumen, $\int_C 1_C$ existe. Para $\varepsilon > 0$, existe una partición P de A tal que $U(1_C, P) - \int_C 1_C < \varepsilon$. Pero $U(1_C, P) = \int_L 1_L$, donde $L = \bigcup\{S_i \in P \mid S_i \cap C \neq \varnothing\} \supseteq C$. De este modo, $v(L \setminus C) = \int_L 1_{L\setminus C} = \int_L (1_L - 1_C) = \int 1_L - \int 1_C < \varepsilon$. Úsese un argumento similar para obtener $K \subset C$.

25. $\int_D 1_D = \int_1^3 \int_{x^2}^{1+x^2} dydx = \int_1^3 dx = 2.$

Capítulo 10: Análisis de Fourier

10.1 Espacios con producto interno

1. **a.** $\langle f, g\rangle = 0 \Rightarrow \langle g, f\rangle = \overline{\langle f, g\rangle} = 0$, por lo que $\|f+g\|^2 = \langle f+g, f+g\rangle = \langle f, f\rangle + \langle f, g\rangle + \langle g, f\rangle + \langle g, g\rangle = \|f\|^2 + \|g\|^2$.

3. Desarróllese $\langle f-g, f\rangle$ usando las propiedades de un producto interno y $\langle \varphi_i, \varphi_j\rangle = \delta_{ij}$.

5. Cualquier sucesión convergente en un espacio con producto interno (o incluso en un espacio métrico general) está acotada.

10.2 Familias de funciones ortogonales

1. Cualquier conjunto ortonormal de vectores es linealmente independiente, ya que si $c_1\varphi_1 + \cdots + c_n\varphi_n = 0$, entonces $c_i = \langle c_1\varphi_1 + \cdots + c_n\varphi_n, \varphi_i\rangle = \langle 0, \varphi_i\rangle = 0$. Esto se cumple para cada i, de modo que la familia es linealmente independiente. Si existen n de ellos en $\mathbb{R}^n$, entonces deben formar una base.

3. **a.** $\langle \psi_n, \psi_m\rangle = \int_0^l \left[\sqrt{2\pi/l}\,\varphi_n(2\pi x/l)\right] \overline{\left[\sqrt{2\pi/l}\,\varphi_n(2\pi x/l)\right]} dx = (2\pi/l)\int_0^l \varphi_n(2\pi x/l)\overline{\varphi_m(2\pi x/l)}\,dx$. Sea $u = 2\pi x/l$, entonces $\langle \psi_n, \psi_m\rangle = (2\pi/l)\int_0^{2\pi} \varphi_n(u)\overline{\varphi_m(u)}(l/2\pi)\,du = \delta_{nm}$.

5. $$\sum_{k=1}^{n}\cos(k\theta) = \mathrm{Re}\left(\sum_{k=1}^{n}(e^{i\theta})^k\right) = \mathrm{Re}\left(e^{i\theta}\frac{1-e^{in\theta}}{1-e^{i\theta}}\right).$$

Muéstrese que esto es

$$\frac{\mathrm{sen}((n+(1/2))\theta)}{\mathrm{sen}(\theta/2)} - \frac{1}{2}.$$

10.3 Completitud y teoremas de convergencia

1. **a.**

$$\begin{aligned} S_N &= \sum_{n=-N}^{N}\left\langle f(x), e^{inx}/\sqrt{2\pi}\right\rangle e^{inx}/2\pi \\ &= \langle f(x), 1\rangle/2\pi + (1/2\pi)\sum_{n=1}^{N}\left(\langle f(x), e^{inx}\rangle e^{inx} + \langle f(x), e^{-inx}\rangle e^{-inx}\right). \end{aligned}$$

Pero

$$\begin{aligned}\left\langle f(x), e^{inx}\right\rangle e^{inx} &= \langle f(x), \cos nx\rangle \cos nx + \langle f(x), \operatorname{sen} nx\rangle \operatorname{sen} nx \\ &+ i(\langle f(x), \cos nx\rangle \operatorname{sen} nx - \langle f(x), \operatorname{sen} nx\rangle \cos nx)\end{aligned}$$

y

$$\begin{aligned}\left\langle f(x), e^{-inx}\right\rangle e^{-inx} &= \langle f(x), \cos nx\rangle \cos nx + \langle f(x), \operatorname{sen} nx\rangle \operatorname{sen} nx \\ &+ i(\langle f(x), \operatorname{sen} nx\rangle \cos nx - \langle f(x), \cos nx\rangle \operatorname{sen} nx).\end{aligned}$$

Así,

$$S_N = \frac{1}{2\pi}\langle f(x), 1\rangle + \frac{1}{\pi}\sum_{n=1}^{N}\left[\langle f(x), \cos nx\rangle \cos nx + \langle f(x), \operatorname{sen} nx\rangle \operatorname{sen} nx\right].$$

b. $\langle f(x), \operatorname{sen} nx\rangle = \int_{-\pi}^{\pi} f(x) \operatorname{sen} nx\, dx = 0$ pues el integrando es una función impar de x.

$$\begin{aligned}\Big(&= \int_{-\pi}^{0} f(x)\operatorname{sen} nx\, dx + \int_{0}^{\pi} f(x)\operatorname{sen} nx\, dx \\ &= \int_{\pi}^{0} f(-x)\operatorname{sen}(-nx)\,(-dx) + \int_{0}^{\pi} f(x)\operatorname{sen} nx\, dx \\ &= \int_{\pi}^{0} f(x)\operatorname{sen}(nx)\, dx + \int_{0}^{\pi} f(x)\operatorname{sen} nx\, dx = 0\Big).\end{aligned}$$

3. Úsese **10.3.3** y cámbiese de variables.

5. **a.** Caso 1: $n \geq 0$ par. Entonces la serie de Fourier converge uniformemente a $f(x)$, pues f es continua y f' es continua a trozos. Caso 2: $n > 0$ impar. Entonces la serie de Fourier converge puntualmente a $f(x)$ para $-\pi < x < \pi$ y a 0 para $x = \pm\pi$. Caso 3: $n \leq -1$. Entonces $f(x) = x^n$ no es ni integrable ni de cuadrado integrable, luego no tiene serie de Fourier.

c. f no es ni integrable ni de cuadrado integrable, luego la serie de Fourier no está definida.

10.4 Funciones de variación acotada y teoría de Fejér

1. $\delta' = \sum_{-\infty}^{\infty}(in/2\pi)e^{inx}$.

10.5 Cálculo de series de Fourier

1. Por la fórmula **2a**, sea $f(x) = 0$ para $-\pi \le x < 0$ y $f(x) = x$ para $0 \le x < \pi$, extendida periódicamente haciendo $f(x + 2\pi) = f(x)$. Entonces $a_0 = (1/\pi) \int_0^\pi x\, dx = \pi/2$. Para $n \neq 0$ tenemos $a_n = (1/\pi) \int_0^\pi x \cos nx\, dx$. Esto se anula si n es par y es igual a $2/\pi n^2$ si n es impar. $b_n = (1/\pi) \int_0^\pi x \operatorname{sen} nx\, dx = (-1)^{n+1}/n$. Así, la serie de Fourier es

$$\frac{\pi}{4} - \frac{2}{\pi} \sum_{k=1}^{\infty} \frac{\cos(2k-1)}{(2k-1)^2} - \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \operatorname{sen} nx.$$

3. La extensión par de $f(x) = x^2$ en $[0, \pi]$ es justamente x^2 en $[-\pi, \pi]$. Así, la serie de cosenos en un semiintervalo es igual a la serie de Fourier de $f(x) = x^2$ en $[-\pi, \pi]$. Ésta y la serie de senos en un semiintervalo están en el apartado **3** de la tabla **10.5-4.**

5. $U = \lim_{n\to\infty} S_n(\pi/2n) = (8 - (-4))(0.089) + 8 \approx 9.068$. Aproximadamente un 13 por ciento de sobreestimación.

$L = \lim_{n\to\infty} S_n(-\pi/2n) = -4 + (8 - (-4))(0.089) \approx -5.068$. Aproximadamente un 27 por ciento de sobreestimación.

10.6 Más teoremas de convergencia

1. La serie de Fourier de f converge absoluta y uniformemente y se puede derivar término a término para obtener la serie de Fourier absoluta y uniformemente convergente de f'.

3. La serie de Fourier converge en media a f y, para $x \neq \pm 1/2$, la serie de Fourier converge a $f(x)$. La serie no se puede derivar término a término en los puntos $\pm 1/2$.

5. Úsese **10.6.2** y el teorema de Parseval.

10.7 Aplicaciones

1. $$y\left(x, \frac{l}{c}\right) = \sum_{n=1}^{\infty} \frac{(-1)^n}{\pi(2n-1)^2} 4h \operatorname{sen}\left(\frac{\pi x}{l}(2n-1)\right),$$

e

$$y\left(x, \frac{3l}{2c}\right) = \frac{1}{2}\left(f\left(x + \frac{3l}{2c}\right) + f\left(x - \frac{3l}{2c}\right)\right) = -f(x).$$

5. $$f(x,\tau) = \frac{l^2}{3} + \frac{4l^2}{\pi^2}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}e^{-n^2\pi^2\tau/l^2}\cos\left(\frac{n\pi x}{l}\right),$$

y $\lim_{\tau\to\infty} f(x,\tau) = l^2/3$.

10.8 Integrales de Fourier

1. Derívese bajo el signo de integral (véase §**9.7**).

3. **a.** Por el ejercicio 1, tenemos $\partial\hat{f}/\partial t = -k^2\alpha^2\,\hat{f}(\alpha,t)$. Intégrese y úsese $\hat{f}(\alpha,0) = \hat{g}(\alpha)$ para obtener $\hat{f}(\alpha,t) = \hat{g}(\alpha)e^{-k^2\alpha^2 t}$.

b. Úsese el teorema enunciado en el texto relativo a las convoluciones para determinar la transformada inversa de Fourier de $\hat{f}(\alpha,t)$, junto con el resultado relativo a la transformada de Fourier de la gaussiana establecido después de la ecuación (10) de §**10.8**.

10.9 Formalismo de la mecánica cuántica

1. $\langle (AB)^*x, y\rangle = \langle x, ABy\rangle = \langle A^*x, By\rangle = \langle B^*A^*x, y\rangle$. Así, $\langle((AB)^* - B^*A^*)x, y\rangle = 0$ para todo par de vectores x e y. Esto implica que $(AB)^* = B^*A^*$ (¿por qué?).

3. $\langle A\psi,\psi\rangle = \langle A\left(\sum_{n=1}^{\infty}\langle\psi,\varphi_n\rangle\,\varphi_n\right), \sum_{k=1}^{\infty}\langle\psi,\varphi_k\rangle\,\varphi_k\rangle =$
$\sum_{n,k}\langle\psi,\varphi_n\rangle\overline{\langle\psi,\varphi_k\rangle}\langle A(\varphi_n),\varphi_k\rangle = \sum_{n,k}\langle\psi,\varphi_n\rangle\overline{\langle\psi,\varphi_k\rangle}\langle\lambda_n\varphi_n,\varphi_k\rangle =$
$\sum_{n,k}\langle\psi,\varphi_n\rangle\overline{\langle\psi,\varphi_k\rangle}\lambda_n\langle\varphi_n,\varphi_k\rangle = \sum_{n=1}^{\infty}\lambda_n|\langle\psi,\varphi_n\rangle|^2$.

Esto funcionaría en un espacio de dimensión finita. De hecho, funciona en general. El valor esperado de A es justamente la suma de los observables $\{\lambda_n\}$ de A ponderados con la probabilidad con la que pueden ser observados cuando A actúa sobre el estado ψ.

5. **a.** Como $\|\psi\| = 1$,

$$\begin{aligned}\Delta^2(A,\psi) &= \langle A^2\psi,\psi\rangle - 2\langle A\psi,\psi\rangle^2 + \langle A\psi,\psi\rangle^2\langle\psi,\psi\rangle\\ &= \langle A^2\psi,\psi\rangle - \langle A\psi,\psi\rangle^2.\end{aligned}$$

c. $$(AB-BA)\psi = A\left(\frac{\hbar}{i}\frac{\partial\psi}{\partial x}\right) - B(x\psi) = \frac{\hbar}{i}x\frac{\partial\psi}{\partial x} - \frac{\hbar}{i}\left(\psi + x\frac{\partial\psi}{\partial x}\right) = \frac{\hbar}{i}\psi,$$

de modo que $|\langle C\psi,\psi\rangle| = \hbar$. Por lo tanto, $\Delta^2(A,\psi)\,\Delta^2(B,\psi) \geq 4\hbar^2$, por el apartado **b**.

Ejercicios del capítulo 10 (al final del capítulo)

1. Si $f_1, f_2 \in M^\perp$, entonces $\langle af_1 + bf_2, g\rangle = a\langle f_1, g\rangle + b\langle f_2, g\rangle = 0$ para toda g en M. Así, $af_1 + bf_2 \in M^\perp$. Por lo tanto, $M^\perp$ es un subespacio. Supóngase que las f_n están en $M^\perp$ y $\|f_n - f\| \to 0$. Sea $g \in M$. Entonces $|\langle f, g\rangle| = |\langle f - f_n, g\rangle + \langle f_n, g\rangle| = |\langle f - f_n, g\rangle| \le \|f - f_n\| \cdot \|g\| \to 0$ cuando $n \to \infty$. Así, $\langle f, g\rangle = 0$ para cada g en M, por lo que $f \in M^\perp$.

5. Sean f_n y f de cuadrado integrable en $[a, b]$ y supongamos que $f_n \to f$ uniformemente en $[a, b]$. Dado $\varepsilon > 0$, elíjase N tal que $|f_n(x) - f(x)| < \varepsilon/\sqrt{b-a}$ para todo x en $[a, b]$. Entonces $\|f - f_n\|^2 = \int_a^b |f(x) - f_n(x)|^2\, dx < \int_a^b \varepsilon^2/(b-a)\, dx = \varepsilon^2$.

7. **a.** $\langle \varphi_i, \varphi_j\rangle = \sum_{k=1}^\infty \delta_{ik}\delta_{kj}$, que es 1 si $i = j$ y 0 si $i \neq j$. Así, las funciones φ_i forman una familia ortonormal. Además, para $x \in l_2$ tenemos $x = \sum_{i=1}^\infty x_i\varphi_i$, de modo que la familia es completa.

9. Las discontinuidades de $\alpha f + \beta g$ se encuentran en la unión de las discontinuidades de f y las discontinuidades de g.

11. Las discontinuidades de $|f|$ están entre las de f.

13. Sea $t_k = \langle f, \varphi_k\rangle$, y obsérvese que $\|f - \sum t_k\varphi_k\|^2 \ge 0$; haga $n \to \infty$.

15. Si $\langle \varphi_n\rangle_1^\infty$ es una familia ortonormal y $n \neq k$, entonces $\|\varphi_n - \varphi_k\| = \sqrt{2}$. No puede existir una subsucesión convergente aunque $\|\phi_n\| = 1$ para cada n.

17. **a.** Sea $s_n = \sum_{k=0}^n c_k\varphi_k$. Entonces $\lim_{n\to\infty} \|s_n\| = 0$. Fijemos k y supóngase que $n > k$. Entonces $|c_k| = |\langle \varphi_k, s_n\rangle| \le \|\varphi_k\| \cdot \|s_n\| = \|s_n\| \to 0$ cuando $n \to \infty$. Luego $c_k = 0$.

19. **a.** Calcúlese

$$\begin{aligned}
(\lambda - \mu)p(x)\,f(x)g(x) &= (\lambda p(x)\,f(x))g(x) - (\mu p(x)g(x))\,f(x)\\
&= [-f''(x) - q(x)\,f(x)]g(x) - [-g''(x) - q(x)g(x)]\,f(x)\\
&= g''(x)\,f(x) - f''(x)g(x)\\
&= [g''(x)\,f(x) + g'(x)\,f'(x)] - [f'(x)g'(x) + f''(x)g(x)]\\
&= [g'(x)\,f(x)]' - [f'(x)g(x)]' = \frac{d}{dx}\left(g'(x)\,f(x) - f'(x)g(x)\right).
\end{aligned}$$

Así, $(\lambda - \mu)\int_a^b p(x)f(x)g(x)\,dx = (\lambda - \mu)[g'(x)f(x) - f'(x)g(x)]\,\Big|_a^b = 0$. Como $\lambda - \mu \neq 0$, la integral debe ser igual a 0.

21. Sea $A = f - s_n$, $B = s_{n+p} - s_n$. Entonces $\langle s_n, B\rangle = 0$ implica $\langle A, B\rangle = \sum_{i=n+1}^{p} \overline{\langle f, \varphi_i\rangle}$ $\langle f, \varphi_i\rangle = \|B\|^2$, y $\langle B, A\rangle = \overline{\langle A, B\rangle} = \|B\|^2$. Así, $\langle f - s_{n+p}, f - s_{n+p}\rangle = \langle A - B, A - B\rangle =$ $\langle A, A\rangle - \langle A, B\rangle - \langle B - A\rangle + \langle B, B\rangle = \|A\|^2 - \|B\|^2 \leq \|A\|^2$. De este modo, $\|f - s_{n+p}\| \leq \|f - s_n\|$.

23.
$$i\hbar\frac{\partial\psi}{\partial t} = i\hbar\left[\sum_{n=0}^{\infty}\langle \psi_0, \varphi_n\rangle\, e^{-iE_nt/h}\left(-\frac{iE_n}{\hbar}\right)\varphi_n\right]$$
$$= \sum_{n=0}^{\infty}\langle \psi_0, \varphi_n\rangle\, e^{-iE_nt/\hbar}E_n\varphi_n = \sum_{n=0}^{\infty}\langle \psi_0, \varphi_n\rangle\, e^{-iE_nt/\hbar}H(\varphi_n)$$
$$= H\left(\sum_{n=0}^{\infty}\langle \psi_0, \varphi_n\rangle\, e^{-iE_nt/\hbar}\varphi_n\right) = H(\psi).$$

Además, $\psi(0, x, y, z) = \sum_{n=0}^{\infty}\langle \psi_0, \varphi_n\rangle\, \varphi_n(x, y, z) = \psi_0$.

25. $\frac{d}{dt}\langle \psi, h(\psi)\rangle = \left\langle \frac{\partial\psi}{\partial t}, H(\psi)\right\rangle + \left\langle H(\psi), \frac{\partial\psi}{\partial t}\right\rangle = \frac{-i}{\hbar}\langle H(\psi), H(\psi)\rangle +$ $\frac{i}{\hbar}\langle H(\psi), H(\psi)\rangle = 0$ ya que H es simétrica.

27. Las autofunciones son $\psi_n(x) = \sqrt{1/l}\,\text{sen}(2\pi nx/l)$ y $\hat{\psi}_n = \sqrt{1/l}\cos((-2n-1)\pi x)/2l$ para $n = 1, 2, \ldots$. Los autovalores son $\varepsilon_n = \hbar^2 2\pi^2 n^2/ml^2$ y $\hat{\varepsilon}_n = \hbar^2(2n-1)^2\pi/2ml^2$.

31. Calcúlese

$$\begin{aligned} f * g(x) &= \int_0^{2\pi} f(y)g(x-y)\,dy = \int_x^{x-2\pi} f(x-w)g(w)(-dw) \\ &= \int_{x-2\pi}^{x} f(x-w)g(w)\,dw = \int_x^{x+2\pi} g(w)f(x-w)\,dw \\ &= \int_0^{2\pi} g(w)\,f(x-w)\,dw \text{ pues el integrando es } 2\pi\text{-periódico.}\end{aligned}$$

Así, $(f * g)(x) = (g * f)(x)$. Además, $(f * g)(x) = \sum_{n=-\infty}^{\infty}(2\pi a_n b_n \sqrt{2\pi})\, e^{inx}/\sqrt{2\pi}$. Luego la relación de Parseval implica $\|f * g\|^2 = 8\pi^3 \sum_{-\infty}^{\infty} |a_n|^2\, |b_n|^2$.

33. Muéstrese que si $|f'(x)| \leq M$ para todo x y $P = \{x_0, x_1, \ldots, x_n\}$ es cualquier partición, entonces $\sum_{j=0}^{n-1} |f(x_{j+1}) - f(x_j)| \leq M(b-a)$.

35. **a.** Puntualmente a $f(x)$ para $x \neq \pm\pi$ y a 0 para $x = \pm\pi$.

c. Puntualmente a $(f(x_0^+) + f(x_0^-))/2$ y en media.

e. $f'(x) = 0$ para $x < 0$ y $2x \operatorname{sen}(1/x) - \cos(1/x)$ para $x > 0$. Así, f' es continua a trozos y está acotada por $2\pi + 1$. Por el ejercicio 33, se puede aplicar el teorema de Jordan-Dirichlet. Así, la convergencia es puntual a $(f(x_0^+) + f(x_0^-))/2$.

37. En $[0, 2\pi]$. $a = 4/\pi$, $b = 0$, $c = 4/3\pi$. En $[-\pi, \pi]$, $a = b = c = 0$.

39. **a.** Si $x_0 < x < x_0 + h$, entonces $\lim_{h\to 0^+} |f(x_0 + h) - f(x)| = \lim_{h\to 0^+} f'(\eta)h = \lim_{\eta\to 0^+} f'(\eta) . \lim_{h\to 0^+} h = 0$. Así, este "criterio de Cauchy" implica que $\lim_{x\to x_0^+} f(x)$ existe.

c. **i.** No existe; no existe.

ii. Existe; no existe.

iii. Ambos existen.

41. Como $\lim_{x\to x_0^+} g'(x) = \lim_{x\to x_0^+} [f'(x)\varphi(x) + f(x)\varphi'(x)] = f'(x_0^+)\varphi(x_0) + f(x_0^+)\varphi'(x_0)$, vemos que $g'(x_0^+)$ existe. Análogamente, $g'(x_0^-)$ existe. Por el ejercicio 39a, $g(x_0^+)$ y $g(x_0^-)$ existen.

43. Aplíquese la desigualdad de Cauchy-Schwarz a las integrales

$$\int_a^b |f(y)|\,|g(y) - g_n(y)|\,dy \qquad \text{y} \qquad \int_a^b |f(y) - f_n(y)|\,|g_n(y)|\,dy.$$

45. **a.** $x \operatorname{sen} x$ es una función par, por lo que $b_n = 0$ y $a_n = (2/\pi)\int_0^\pi x \operatorname{sen} x \cos nx\,dx = (1/\pi)\int_0^\pi x\,[\operatorname{sen}(n+1)x - \operatorname{sen}(n-1)x]\,dx$. Para $n = 0$, esto es igual a 2; para $n = 1$ es $-1/2$, y para $n \geq 2$ es $(-1)^{n+1}\, 2/(n^2 - 1)$. Así, la serie de Fourier es $2/2 - (1/2)\cos x - 2\sum_{n=2}^{\infty} ((-1)^n/(n^2-1)) \cos nx$.

47. $\lim_{n\to\infty} s_n(\pi/2n) \approx (2-0)(0.089) = 2.178$. $\lim_{n\to\infty} s_n(-\pi/2n) \approx 0 - (2-0)(0.089) = -0.178$.

51. Las filas de la tabla coinciden con las de la tabla 10.5-4.

Tabla del ejercicio 51			
Fila de la tabla	Discontinuidades	En el interior	Límite de la serie
1	$0, \pm\pi$	1/2	1/2 en $\pm\pi$
1*a*	$0, \pi$	0	0 en $\pm\pi$
	ninguno	—	1 en 0 y 1
2	$\pm\pi$	—	0 en $\pm\pi$
	$0, 2\pi$	—	π en $0, 2\pi$
	ninguno	—	0 en 0, π en π
2*a*	$\pm\pi$	—	$\pi/2$ en $\pm\pi$
3	$0, 2\pi$	—	$2\pi^2$ en $0, 2\pi$
	ninguno	—	π^2 en $\pm\pi$
	$\pm\pi$	—	0 en $0, \pi$
4	ninguno	—	0 en $0, \pi$
4*a*	ninguno	—	0 en $0, \pm\pi$
4*b*	ninguno	—	—
5	$0, \pi$	—	0 en $0, \pi$
6	$\pm\pi$	—	senh (π) en $\pm\pi$

55. **a.** Sea $h(x) = f(x)$ si $0 \leq x \leq \pi$ y $f(-x)$ si $-\pi \leq x < 0$. Entonces $h(x)$ es continua en $[-\pi, \pi]$ y h' es continua a trozos, con discontinuidades de salto. Aplíquese el teorema de la convergencia uniforme, **10.6.1.**

59. Aplíquese la desigualdad de Cauchy-Schwarz a las sumas.

61. Dada una partición P de $[0, 2\pi]$, calcúlese

$$\begin{aligned}\sum_{i=1}^{n} |g(x_i) - g(x_{i-1})| &= \sum_{i=1}^{n} \left| \int_{x_{i-1}}^{x_i} f(y)\, dy \right| \\ &\leq \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} |f(y)|\, dy \leq \int_0^{2\pi} |f(y)|\, dy,\end{aligned}$$

que es finita por la desigualdad de Schwarz y porque f es de cuadrado integrable.

63. **a.** En media y puntualmente, excepto para $x = \pm\pi$. No se puede derivar término a término.

c. En media, puntualmente y uniformemente. Se puede derivar.

e. Lo mismo que en **a.**

g. En media y puntualmente, no uniformemente. No se puede derivar.

67. $$\varphi(x, y) = \sum_{n=1}^{\infty} \left[\frac{\pi((-1)^n - 2)}{2n} + \frac{(1 - (-1)^n)}{\pi n^3} \right] \operatorname{sen} nx \frac{\operatorname{senh}(n(\pi - y))}{\operatorname{senh}(n\pi)}.$$

b. $\varphi(x,y) = 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \text{sen}(nx) \frac{\text{senh}(n(\pi - y))}{\text{senh}(n\pi)}.$

71. **b.**
$$f(x) = \frac{\tilde{a}_0}{2k} + \sum_{n=1}^{\infty} \left(k - \frac{n^2\pi^2}{l^2}\right)^{-1} \left(\tilde{a}_n \cos\left(\frac{n\pi x}{l}\right) + \tilde{b}_n \,\text{sen}\left(\frac{n\pi x}{l}\right)\right).$$

73. Usando el lema de Riemann-Lebesgue:

$$\begin{aligned}\lim_{k\to\infty} \int_0^{\pi} \sqrt{x}\,\text{sen}^2(kx)\,dx &= (1/2)\int_0^{\pi} \sqrt{x}\,dx \\ &\quad -(1/2)\lim_{k\to\infty}\int_0^{\pi} \sqrt{x}\cos(2kx)\,dx \\ &= (1/3)\pi^{3/2} - 0 = (1/3)\pi^{3/2}.\end{aligned}$$

Índice de materias